

# BREVET D'INVENTION

CERTIFICAT D'UTILITÉ - CERTIFICAT D'ADDITION

**PRIORITY DOCUMENT**

**COPIE OFFICIELLE**

Le Directeur général de l'Institut national de la propriété industrielle certifie que le document ci-annexé est la copie certifiée conforme d'une demande de titre de propriété industrielle déposée à l'Institut.

Fait à Paris, le 24 JAN. 1996

Pour le Directeur général de l'Institut  
national de la propriété industrielle  
Le Chef de Division

Yves CAMPENON

INSTITUT  
NATIONAL DE  
LA PROPRIÉTÉ  
INDUSTRIELLE

SIEGE  
26 bis, rue de Saint Petersburg  
75800 PARIS Cédex 08  
Téléphone : (1) 42 94 52 52  
Télécopie : (1) 42 93 59 30



# REQUETE

## EN DÉLIVRANCE D'UN TITRE DE PROPRIÉTÉ INDUSTRIELLE \*

1

a	<input checked="" type="checkbox"/>	BREVET D'INVENTION
b	<input type="checkbox"/>	CERTIFICAT D'UTILITÉ
c	<input type="checkbox"/>	DEMANDE DIVISIONNAIRE
d	<input type="checkbox"/>	TRANSFORMATION D'UNE DEMANDE DE BREVET EUROPEEN

Pour c et d, précisez : Nature, N° et date de la demande initiale

2 OPTIONS OBLIGATOIRES au moment du dépôt (sauf pour le certificat d'utilité)

LE DEMANDEUR REQUIERT L'ETABLISSEMENT DIFFERE DU RAPPORT DE RECHERCHE

☐ OUI ☒ NON

SI L'OPTION CHOISIE EST NON ET SI LE DEMANDEUR EST UNE PERSONNE PHYSIQUE IL REQUIERT LE PAIEMENT ECHELONNE DE LA REDEVANCE DE RAPPORT DE RECHERCHE

☐ OUI ☒ NON

NATURE NUMERO DATE DE LA DEMANDE INITIALE

DATE DE REMISE DES PIECES

31 MAR 1995

N° D'ENREGISTREMENT NATIONAL

95 03841 -

CODE POSTAL DU LIEU DE DEPOT

75

DATE DE DEPOT

31 MARS 1995

4 NUMERO DU POUVOIR PERMANENT

15 Janvier 1991

3 NOM ET ADRESSE DU DEMANDEUR OU DU MANDATAIRE A CUI TOUTE LA CORRESPONDANCE DOIT ETRE ADRESSEE

**RHONE-POULENC RORER S.A.**  
Direction Brevets  
20, Avenue Raymond Aron  
92160 ANTONY  
FRANCE

5 REFERENCE DU CORRESPONDANT

PB ST 95021

6 TELEPHONE DU CORRESPONDANT

40 91 70 36

7 TITRE DE L'INVENTION

SYSTEME D'EXPRESSION CONDITIONNEL

8 DEMANDEUR(S) : Nom et Prénoms (souligner le nom patronymique) ou dénomination et forme juridique

N° SIREN.

3 0 4 4 6 3 2 8 4

**RHONE-POULENC RORER S.A.**

9 ADRESSE(S) COMPLÈTE(S)

20, Avenue raymond Aron  
92160 ANTONY

PAYS

FRANCE

10 NATIONALITÉ(S)

Française

11 INVENTEUR(S)

LE DEMANDEUR EST L'UNIQUE INVENTEUR

☐ OUI ☒ NON

Si la réponse est non voir notice explicative

12

SI LE DEMANDEUR EST UNE PERSONNE PHYSIQUE NON IMPOSABLE, IL REQUIERT QU'A REQUIS LA REDUCTION DES REDEVANCES

☐ OUI ☒ NON

REDEVANCES VERSÉES

☒ DE DÉPÔT

☒ DE RAPPORT DE RECHERCHE

☐ DE REVENDICATION DE PRIORITÉ

☒ DE REVENDICATION (à partir de la 11e)

13 DECLARATION DE PRIORITÉ  
OU REQUÊTE DU BÉNÉFICE DE  
LA DATE DE DÉPÔT D'UNE  
DEMANDE ANTERIEURE

PAYS D'ORIGINE

DATE DE DEPOT

NUMERO

14

VISIONS

ANTERIEURES A LA  
PRESENTE DEMANDE

N°

N°

N°

N°

15 SIGNATURE DU DEMANDEUR OU DU MANDATAIRE  
NOM ET QUALITÉ DU SIGNATAIRE N° D'INSCRIPTION

SIGNATURE DU PREPOSE A LA RECEPTION

SIGNATURE APRES ENREGISTREMENT DE LA DEMANDE A L'INPI

P. BECKER

*[Signature]*

Division Administrative des Brevets

**DÉSIGNATION DE L'INVENTEUR**

(si le demandeur n'est pas l'inventeur ou l'unique inventeur)

N° d'enregistrement national

9503841

**Titre de l'invention :**

**SYSTEME D'EXPRESSION CONDITIONNEL**

**Le (s) soussigné (s)**

**RHONE-POULENC RORER S.A.**  
20, Avenue Raymond Aron  
92160 ANTONY  
FRANCE

**désigne (nt) en tant qu'inventeur (s) (indiquer nom, prénoms, adresse et souligner le nom patronymique) :**

BRACCO Laurent - 12, rue Moulin des prés, 75013 PARIS - FRANCE

SCHWEIGHOFFER Fabien - 53, Bld de la Libération, 94300 VINCENNES - FRANCE

TOCQUE Bruno - 58, Bld Saint-Denis, 92400 COURBEVOIE - FRANCE

**NOTA : A titre exceptionnel, le nom de l'inventeur peut être suivi de celui de la société à laquelle il appartient (société d'appartenance) lorsque celle-ci est différente de la société déposante ou titulaire.**

Date et signature (s) du (des) demandeur (s) ou du mandataire  
Antony, le 31 Mars 1995



Philippe BECKER

La présente invention concerne un nouveau système pour l'expression conditionnelle de gènes. Elle concerne également l'utilisation de ce système en thérapie génique ou cellulaire, pour cibler de manière très sélective l'expression de gènes d'intérêt.

Les thérapies génique et cellulaire consistent à corriger une déficience ou une anomalie (mutation, expression aberrante, etc) ou à assurer l'expression d'une protéine d'intérêt thérapeutique par introduction d'une information génétique dans la cellule ou l'organe affecté. Cette information génétique peut être introduite soit ex vivo dans une cellule extraite de l'organe, la cellule modifiée étant alors réintroduite dans l'organisme (thérapie cellulaire), soit directement in vivo dans le tissu approprié (thérapie génique). Différentes techniques existent pour effectuer le transfert de gènes, parmi lesquelles des techniques diverses de transfection impliquant des vecteurs chimiques ou biochimiques, naturels ou synthétiques tels que des complexes d'ADN et de DEAE-dextran (Pagano et al., J.Virol. 1 (1967) 891), d'ADN et de protéines nucléaires (Kaneda et al., Science 243 (1989) 375), d'ADN et de lipides (Felgner et al., PNAS 84 (1987) 7413), l'emploi de liposomes (Fraley et al., J.Biol.Chem. 255 (1980) 10431), lipides cationiques, etc. Une autre technique repose sur l'emploi de virus comme vecteurs pour le transfert de gènes. A cet égard, différents virus ont été testés pour leur capacité à infecter certaines populations cellulaires. En particulier, les rétrovirus (RSV, HMS, MMS, etc), le virus HSV, les virus adéno-associés et les adénovirus. L'une des difficultés majeures pour le développement de ces thérapies génique et cellulaire réside cependant dans la sélectivité du traitement. Selon les applications, selon le gène à transférer, il est important de pouvoir cibler certains tissus ou certaines parties seulement de l'organisme afin de concentrer l'effet thérapeutique et de limiter la dissémination et les effets secondaires. Ce ciblage peut être réalisé en utilisant des vecteurs présentant une spécificité cellulaire donnée. Une autre approche consiste à utiliser des signaux d'expression spécifiques de certain types cellulaires. A cet égard, des promoteurs dits spécifiques ont été décrits dans la littérature, tels que le promoteur des gènes codant pour la pyruvate kinase, la villine, la GFAP, le promoteur de la protéine intestinale de liaison

des acides gras, le promoteur de l'actine  $\alpha$  des cellules du muscle lisse, ou le promoteur des gènes de l'apo-AI, l'apo-AII, l'albumine humaine, etc. Cependant, ces promoteurs présentent certains inconvénients et en particulier ils présentent un certain bruit de fond transcriptionnel qui peut être  
5 gérant pour l'expression de gènes toxiques et ils sont limités à certaines cellules et ne peuvent donc être utilisés pour toute application.

La présente invention décrit maintenant un nouveau système d'expression conditionnel de gènes, particulièrement sélectif et efficace. Une des caractéristiques avantageuses du système de l'invention réside dans sa  
10 capacité à exprimer un gène non pas en fonction d'un type cellulaire, mais en fonction de la présence d'un élément cellulaire particulier ou d'une situation physiologique particulière. Ce système fait en effet appel à des molécules chimériques bispécifiques comprenant un domaine capable de lier sélectivement une séquence définie d'ADN et un domaine détecteur capable  
15 de lier spécifiquement un transactivateur ou un complexe transactivateur.

Un premier aspect de la présente invention réside plus particulièrement dans la création et l'expression de molécules chimériques bispécifiques comprenant un domaine capable de lier sélectivement une séquence définie d'ADN et un domaine capable de lier spécifiquement un  
20 transactivateur ou transrépresseur ou un complexe transactivateur ou transrépresseur.

Un autre aspect de la présente invention réside dans une séquence d'acide nucléique codant pour une molécule chimérique telle que définie ci-avant, ainsi que tout un vecteur d'expression comprenant ladite séquence  
25 nucléique.

Un autre aspect de l'invention consiste en un système conditionnel d'expression de gènes comprenant (i) une molécule chimérique telle que définie ci-avant et (ii) une cassette d'expression comportant une séquence de régulation, un promoteur minimal (dont l'activité dépend de la présence d'un  
30 transactivateur) et ledit gène.

Un autre aspect de l'invention réside également dans un vecteur d'expression comprenant

- une séquence nucléique codant pour une molécule chimérique telle que définie ci-avant et
- ladite cassette d'expression.

5 Le système d'expression conditionnel de l'invention est particulièrement approprié pour une utilisation en thérapie génique ou cellulaire, pour cibler de manière très sélective l'expression de gènes d'intérêt.

10 L'une des composantes du système de l'invention consiste donc en des molécules chimériques bispécifiques particulières comprenant un domaine capable de lier sélectivement une séquence définie d'ADN et un domaine capable de lier spécifiquement un transactivateur ou un complexe transactivateur. La bispécificité des molécules de l'invention réside d'une part dans leur capacité à lier une séquence d'ADN définie (généralement désignée séquence régulatrice ou opérateur) et d'autre part dans leur  
15 capacité à recruter spécifiquement un domaine protéique transactivateur ou transrépresseur permettant d'induire ou de réprimer l'expression de gènes.

L'invention porte particulièrement sur la mise au point de molécules chimériques bispécifiques permettant le recrutement de tout facteur transcriptionnel dont l'activation ou l'inactivation conduit à une situation  
20 physiopathologique. Les molécules chimériques bispécifiques selon l'invention permettent ainsi le recrutement sélectif de transactivateurs transcriptionnels spécifiques d'un état physiopathologique, la fixation de ces facteur transcriptionnels à des promoteurs par l'intermédiaire d'une fixation de ces molécules à des séquences d'ADN définies localisées près de ces  
25 promoteurs (séquences régulatrices ou opérateur), et ainsi l'expression conditionnelle de gènes (Figure 1).

L'invention porte également sur la mise au point de molécules chimériques bispécifiques permettant le recrutement non pas d'une molécule portant un domaine transactivateur mais d'un complexe transactivateur  
30 transcriptionnel, c'est à dire d'un complexe formé entre une molécule cible présente dans une cellule et une molécule portant un domaine transactivateur (Figure 2). Dans ce cas, le complexe transactivateur est préférentiellement formé au moyen d'une deuxième molécule chimérique

bispécifique comprenant un domaine transactivateur et un domaine de liaison sélectif à ladite molécule cellulaire. La fixation de cette deuxième molécule permet la formation d'un complexe binaire transactivateur transcriptionnel, lequel complexe étant alors recruté par le système détecteur de l'invention.

- 5 La fixation de ce complexe ternaire à proximité de promoteurs permet ainsi l'expression régulée de gènes. Ce type de construction permet avantageusement d'élargir les conditions d'utilisation du système de l'invention à la détection de toute molécule intracellulaire dépourvue de domaine transactivateur, qu'il s'agisse d'une molécule endogène ou d'une
- 10 molécule d'origine infectieuse par exemple.

- Le système de l'invention permet ainsi, grace à un système de détection très sélectif ("sensor") d'activer l'expression de gènes d'intérêt uniquement en présence de protéines cibles. Il peut s'agir de facteurs transcriptionnels apparaissant lors de situations physiologiques ou
- 15 physiopathologiques, ou de toute molécule endogène ou d'origine infectieuse par exemple. Le système de l'invention contient en effet un élément détecteur très sensible et très sélectif permettant de conditionner l'expression d'un gène à la présence, l'apparition ou la disparition de toute molécule dans une cellule.

- 20 Dans le cadre de la présente invention, le terme transactivateur désigne tout facteur transactivateur de la transcription ou toute protéine comportant un domaine transactivateur transcriptionnel. Le complexe transactivateur désigne le complexe formé entre une molécule présente dans une cellule et une molécule bispécifique de l'invention comprenant un
- 25 domaine transactivateur et un domaine de liaison spécifique à ladite molécule. Le système d'expression de l'invention peut être utilisé pour recruter toute protéine transactivatrice ou portant un domaine transactivateur, et en particulier toute protéine d'origine virale, parasitaire, mycobactérienne ou cellulaire possédant une activité transactivateur
- 30 transcriptionnel. Parmi les facteurs transcriptionnels d'origine virale on peut citer notamment la protéine Tat du virus VIH, les protéines E6/E7 du virus du papillome ou encore la protéine EBNA du virus d'Epstein Barr. Ces protéines possèdent un domaine transactivateur transcriptionnel et sont présentes uniquement dans les cellules infectées par ces virus, c'est-à-dire dans des
- 35 conditions physiopathologiques particulières. Le système d'expression



conditionnel selon l'invention permet avantageusement une détection de cette situation physiologique (l'apparition de ces transactivateurs spécifiques de l'infection virale) et l'induction d'une expression sélective de gène(s) donné(s). Parmi les protéines cellulaires, on peut citer préférentiellement la

5 protéine p53 mutée ou sauvage. La protéine p53 est constituée de 393 acides aminés. Sous sa forme sauvage, elle est un suppresseur de tumeur capable de réguler négativement la croissance et la division cellulaire. Cette activité est liée à la présence d'un domaine transactivateur transcriptionnel dans la structure de la protéine p53, localisé dans la région N-terminale de la

10 protéine (résidus 1-100 environ). Dans certaines situations, p53 sauvage est également capable d'induire l'apoptose (Yonish-Rouach et al., *Nature*, 352, 345-347, 1991). Ces propriétés se manifestant en situation de stress où l'intégrité de l'ADN cellulaire est menacée, p53 a été suggérée être un "gardien du génome". La présence de p53 mutées dans environ 40 % des

15 tumeurs humaines, tous types confondus, renforce cette hypothèse et souligne le rôle probablement crucial que jouent les mutations de ce gène dans le développement tumoral (pour revues, voir Montenarh, *Oncogene*, 7, 1673-1680, 1992 ; Oren, *FASEB J.*, 6, 3169-3176, 1992 ; Zambetti and Levine, *FASEB J.*, 7, 855-865, 1993). Dans le cadre de la présente

20 invention, il est possible de recruter sélectivement le domaine transactivateur de la protéine p53 et ainsi d'induire l'expression contrôlée de gène(s) uniquement dans les cellules contenant cette protéine. Il est particulièrement intéressant selon l'invention de recruter spécifiquement les formes mutées de la protéine p53 qui, comme indiqué ci-avant, apparaissent dans des

25 situations physiopathologiques d'hyperprolifération cellulaire (type cancer). Ce ciblage peut être réalisé préférentiellement au moyen d'un domaine de liaison spécifique aux formes mutées de la protéine p53. Il existe toutefois une spécificité de fait liée à l'accumulation des formes mutées qui possèdent une demi-vie très supérieure à la forme sauvage.

30 Le système de l'invention peut également être utilisé pour induire l'expression sélective de gène(s) par détection de toute molécule cible présente dans une cellule. La protéine détectée est préférentiellement une protéine apparaissant dans une cellule dans des situations anormales (infection, hyperprolifération, etc). Il peut s'agir en particulier de protéines

35 virales telles que des protéines de structure ou de fonction d'un virus, et

notamment du virus VIH, hépatite, herpès, etc. Il peut s'agir également de protéines spécifiques d'un état d'hyperprolifération cellulaire telles que notamment les protéines myc, fos, jun, les cyclines, etc.

5 L'une des propriétés des molécules chimériques de l'invention réside donc dans leur capacité à se lier à des régions spécifiques d'ADN (régions régulatrices ou opérateur). Cette liaison permet d'apporter le domaine transactivateur à proximité du promoteur et de ce fait d'activer l'expression d'un gène placé sous contrôle dudit promoteur.

10 Le domaine capable de lier sélectivement une séquence définie d'ADN présent dans les molécules de l'invention est essentiellement d'origine protéique. Plus préférentiellement, ce domaine dérive d'une protéine procaryote ou eucaryote capable d'interagir avec des séquences d'ADN. De nombreuses études génétiques et structurales ont aujourd'hui permis de  
15 définir précisément, au sein de protéines interagissant avec des séquences d'ADN double brin, les domaines responsables de ces interactions.

Parmi les protéines procaryotes interagissant avec des séquences d'ADN double brin, on peut citer notamment des represseurs bactériens et, préférentiellement, le répresseur tétracycline d'E. Coli et le répresseur Cro du  
20 bactériophage Lambda.

Le répresseur tétracycline (tetR) de E.coli est une protéine de 210 acides aminés environ. Chez E.coli, tetR contrôle négativement la transcription de gènes médiant la résistance à cet antibiotique au sein de l'opéron tet. En absence de tétracycline, le répresseur tetR se fixe à l'ADN au  
25 niveau d'une séquence spécifique (désignée séquence opérateur ou Tetop) et réprime la transcription du gène de résistance. Au contraire, en présence de tétracycline, le répresseur tetR ne se fixe plus à l'opérateur tetop permettant une transcription constitutive du gène (Hillen.W and Wissmann.A. (1989) in Protein-Nucleic Acid Interaction. Topics in Molecular and Structural  
30 Biology.eds, Saenger.W. and Heinemann.U. (Macmillan, London), Vol.10. pp. 143-162). La séquence de tetR a été publiée (elle est reproduite notamment dans WO94/04682). La séquence d'ADN double brin spécifique de liaison de tetR à l'ADN (Tetop) est composée du motif suivant :

TCTCTATCACTGATAGGGA (SEQ ID n° 1).

Ce motif peut être répété plusieurs fois pour augmenter l'affinité et l'efficacité du système. Ainsi, La séquence régulatrice peut comprendre jusqu'à 10 motifs et, de préférence, contient 2 motifs (Tetop2) ou 7 motifs (Tetop7) (voir Figure 3).

5 La protéine Cro a été définie à l'origine comme un régulateur de l'expression du répresseur CI (Eisen.H. et coll. (1970) PNAS 66, pp855). Le clonage du gène *cro* a permis l'identification d'une protéine de 66 acides aminés (SEQ ID n° 21; Roberts;T et coll. (1977) Nature 270, pp274). Cro exerce son rôle physiologique en se fixant préférentiellement à l'opérateur  
10 OR3 de Lambda.

La séquence d'ADN double brin spécifique de liaison de Cro à l'ADN (région désignée OR3) est composée des bases suivantes :

TATCACCGCAAGGGATA (SEQ ID n° 2)

15 Cette région peut également être répétée plusieurs fois pour augmenter l'affinité et l'efficacité du système (voir Figure 4).

Parmi les protéines eucaryotes interagissant avec des séquences d'ADN double brin, on préfère utiliser pour la construction des molécules de l'invention les protéines ou domaines dérivés des protéines STAT, p53 ou NFkB (Inoue et al., PNAS 89 (1992) 4333). Concernant la protéine p53, son  
20 domaine de liaison à l'ADN est localisé dans la région centrale de la protéine et, plus précisément, dans la région comprise entre les acides aminés 102 à 292 (Pavletich et al., Genes & Dev. 7 (1993) 2556).

Comme indiqué ci-avant, le domaine capable de lier sélectivement une  
25 séquence définie d'ADN présent dans les molécules de l'invention est préférentiellement dérivé d'une protéine procaryote ou eucaryote capable d'interagir avec une région d'ADN double-brin. Le domaine utilisé pour la construction des molécules de l'invention peut être constitué de l'ensemble de la protéine ou d'un fragment de celle-ci comportant la région d'interaction  
30 avec l'ADN. Ce domaine a été identifié pour différentes protéines et notamment TetR (voir par exemple Berens et al., J. Biol. Chem. 267 (1992) 1945). Il peut également être constitué d'un dérivé de cette protéine ou du fragment ayant conservé les propriétés de liaison à l'ADN. De tels dérivés  
35 sont notamment des protéines présentant des modifications d'un ou plusieurs acides aminés, par exemple pour permettre leur fusion avec les autres

domaines des molécules de l'invention, préparées selon les techniques classiques de la biologie moléculaire. Des dérivés des protéines TetR et Cro par exemple ont été décrits dans la littérature, possédant des mutations ponctuelles et/ou des délétions (Hecht et al., J. Bact. 175 (1993) p. 1206; 5 Altschmied et al., EMBO J 7 (1988) 4011; Baumeister et al., Proteins 14 (1992) 168; Hansen et al., J. Biol. Chem. 262 (1987) 14030). La capacité de liaison de ces dérivés à une séquence définie d'ADN peut ensuite être testée par incubation du dérivé préparé avec la séquence régulatrice et détection des complexes formés. En outre, les dérivés peuvent également être des 10 protéines ayant des propriétés améliorées de liaison à l'ADN (spécificité, affinité, etc).

Selon un mode préféré de mise en oeuvre, le domaine capable de lier sélectivement une séquence définie d'ADN présent dans les molécules de 15 l'invention est dérivé d'une protéine procaryote. Ce type de construction est particulièrement avantageux puisque ces protéines, d'origine non-humaine, reconnaissent des sites d'ADN double brin d'au moins 14 nucléotides. La probabilité de trouver la même séquence au sein du genome humain est quasi nulle et donc le système d'expression obtenu est d'autant plus sélectif.

20 Dans un mode de réalisation préféré, le domaine capable de lier sélectivement une séquence définie d'ADN présent dans les molécules de l'invention est dérivé des protéines tetR ou Cro. Il est tout particulièrement avantageux d'utiliser les protéines tetR ou Cro (SEQ ID n° 21) complètes.

25 Le domaine capable de lier spécifiquement le transactivateur transcriptionnel ou le complexe transactivateur transcriptionnel présent dans les molécules de l'invention peut être de différents types. Il peut s'agir en particulier d'un domaine d'oligomérisation dans le cas où le transactivateur ou le complexe transactivateur ciblé comporte également un tel domaine. Il peut également s'agir de tout domaine synthétique ou naturel connu pour 30 interagir avec ledit transactivateur ou complexe transactivateur. Il peut encore s'agir d'un anticorps ou d'un fragment ou dérivé d'un anticorps dirigé contre le transactivateur ou complexe transactivateur.

Parmi les domaines d'oligomérisation utilisables dans le cadre de l'invention on peut citer plus particulièrement les leucine-zipper, les domaines

SH2 ou les domaines SH3 par exemple. Les leucine-zipper sont des hélices  $\alpha$  amphipatiques qui contiennent 4 ou 5 leucines tous les 7 acides aminés. Cette périodicité permet la localisation des leucines à peu près à la même position sur l'hélice  $\alpha$ . La dimérisation est sous tendue par des interactions

5 hydrophobes entre les chaînes latérales des leucines de deux domaines zipper contigus (Vogt et al., Trends In Bioch. Science 14 (1989) 172). Les domaines SH2 sont connus pour interagir avec des séquences peptidiques spécifiques phosphorylées en tyrosine. Les domaines SH3 peuvent être

10 utilisés pour former un oligomère avec tout transactivateur ou complexe transactivateur comportant le peptide riche en proline correspondant (Pawson et al., Current Biology 3 (1993) 434). On peut également utiliser des régions de protéines connues pour induire une oligomérisation, telles que notamment la région C-terminale de la protéine p53. L'utilisation de cette

15 région permet de recruter de manière sélective les protéines p53 présentes dans une cellule. On utilise préférentiellement dans le cadre de l'invention une région de p53 comprise entre les acides aminés 320-393 (SEQ ID n° 3), 302-360 ou 302-390.

Parmi les domaines synthétiques ou naturels connus pour interagir avec la molécule comportant l'élément transactivateur ciblé, on peut citer par

20 exemple la région de la protéine MDM2 interagissant avec la protéine p53. Ce type de construction permet ainsi de recruter comme transactivateur la protéine p53 sauvage ou mutée.

Un domaine de liaison spécifique au transactivateur transcriptionnel préféré de l'invention est constitué par un anticorps ou un fragment ou dérivé

25 d'anticorps. Les fragments ou dérivés d'anticorps sont par exemple les fragments Fab ou F(ab)'2, les régions VH ou VL d'un anticorps ou encore des anticorps simple chaîne (ScFv) comprenant une région VH liée à une région VL par un bras. Ce type de domaine est particulièrement avantageux puisqu'il peut être dirigé contre toute molécule.

30 Les anticorps, molécules de la superfamille des immunoglobulines, sont constitués de différentes chaînes (2 lourdes (H) et 2 légères (L)) elles mêmes composées de différents domaines (domaine variable (V) domaine de jonction (J), etc). Le domaine de liaison au transactivateur ou complexe transactivateur présent dans les molécules de l'invention est

avantageusement constitué d'un fragment ou dérivé d'anticorps comprenant au moins le site de liaison de l'antigène. Ce fragment peut être soit le domaine variable d'une chaîne légère ( $V_L$ ) ou lourde ( $V_H$ ), éventuellement sous forme de fragment Fab ou  $F(ab')_2$  ou, préférentiellement, sous forme d'anticorps simple chaîne (ScFv). Les anticorps simple chaîne utilisés pour la construction des molécules de l'invention sont constitués d'un peptide correspondant au site de liaison de la région variable de la chaîne légère d'un anticorps relié par un bras peptidique à un peptide correspondant au site de liaison de la région variable de la chaîne lourde d'un anticorps. La construction de séquences d'acides nucléiques codant pour de tels anticorps modifiés selon l'invention a été décrite par exemple dans le brevet US4,946,778 ou dans les demandes WO94/02610, WO94/29446. Elle est illustrée dans les exemples.

Une construction préférée selon la présente invention comprend un domaine de liaison à une protéine p53. Il s'agit plus préférentiellement d'un dérivé d'anticorps dirigé contre une protéine p53. Un mode de réalisation particulier est constitué par un anticorps simple chaîne dirigé contre p53. A titre d'exemple particulier, on utilise le ScFv de séquence SEQ ID n° 4 dont la construction est décrite dans les exemples.

Le domaine de liaison à l'ADN et le domaine de liaison au transactivateur sont généralement reliés entre eux par l'intermédiaire d'un bras. Ce bras est généralement constitué d'un peptide conférant une flexibilité suffisante pour que les deux domaines des molécules de l'invention puissent être fonctionnels de manière autonome. Ce peptide est généralement composé d'acides aminés non chargés, n'interférant pas avec l'activité des molécules de l'invention, tels que par exemple la glycine, la serine, le tryptophane, la lysine ou la proline. Le bras comprend généralement de 5 à 30 acides aminés et, de préférence, de 5 à 20 acides aminés. Des exemples de bras peptidiques utilisables pour la construction des molécules de l'invention sont par exemple :

- GGGGSGGGGSGGGGS (SEQ ID n° 5)
- PKPSTPPGSS (SEQ ID n° 6) dont la séquence codante est CCCAAGCCCAGTACCCCCCAGGTTCTTCA (SEQ ID n° 6).

Des exemples préférés de molécule selon l'invention sont notamment les molécules suivantes :

a) ScFv-tag-Hinge-TET ou Cro (Figure 5A)

Ce type de molécule comporte :

5       - un domaine de liaison à un transactivateur constitué d'un anticorps simple chaîne,

- une séquence peptidique tag reconnue par un anticorps monoclonal permettant la détection immunologique de la molécule. Cette séquence peut être par exemple l'épitope VSV de séquence MNRLGK (SEQ ID n° 7) dont la  
10       séquence codante est ATGAACCGGCTGGGCAAG (SEQ ID n° 7), ou l'épitope myc de séquence EQKLISEEDLN (SEQ ID n° 8) dont la séquence codante est GAACAAAACTCATCTCAGAAGAGGATCTGAAT (SEQ ID n° 8), reconnu par l'anticorps 9E10.

- un bras peptidique de séquence SEQ ID n° 6 (Hinge) et

15       - un domaine de liaison à l'ADN constitué de la protéine TET ou Cro. Préférentiellement, le ScFv est dirigé contre une protéine p53.

b) ScFv-Hinge-TET ou Cro (Figure 5B)

Ce type de molécule comporte les mêmes éléments que la molécule a) à l'exception de la séquence tag qui est absente.

20       c) ScFv-TET ou Cro (Figure 5C)

Ce type de molécule comporte simplement un domaine de liaison à un transactivateur constitué d'un anticorps simple chaîne et un domaine de liaison à l'ADN constitué de la protéine TET ou Cro. Il ne comporte ni bras, ni séquence tag. Dans cette construction, le domaine de liaison au  
25       transactivateur est localisé dans la partie N-terminale de la molécule et le domaine de liaison à l'ADN dans la partie C-terminale.

d) TET ou Cro-ScFv (Figure 5D)

Ce type de molécule est similaire au type c) ci-dessus. La différence réside essentiellement dans l'agencement des domaines : le domaine de  
30       liaison au transactivateur est maintenant localisé dans la partie C-terminale de la molécule et le domaine de liaison à l'ADN dans la partie N-terminale.

e) TET ou Cro-Hinge-ScFv (Figure 5E)

Ce type de molécule comporte les mêmes éléments que la molécule b) ci-dessus. La différence réside essentiellement dans l'agencement des  
35       domaines : le domaine de liaison au transactivateur est maintenant localisé

dans la partie C-terminale de la molécule et le domaine de liaison à l'ADN dans la partie N-terminale.

f) Oligom-tag-Hinge-TET ou Cro (Figure 5A)

5 Ce type de molécule est semblable au type a), à l'exception du domaine de liaison au transactivateur qui est remplacé par le domaine d'oligomérisation à la protéine p53 de séquence SEQ ID n° 3. Cette molécule permet de recruter les protéines p53 mutées apparaissant dans les cellules tumorales.

g) Oligom-Hinge-TET ou Cro (Figure 5B)

10 Ce type de molécule est semblable au type b), à l'exception du domaine de liaison au transactivateur qui est remplacé par le domaine d'oligomérisation à la protéine p53 de séquence SEQ ID n° 3.

h) Oligom-TET ou Cro (Figure 5C)

15 Ce type de molécule est semblable au type c), à l'exception du domaine de liaison au transactivateur qui est remplacé par le domaine d'oligomérisation à la protéine p53 de séquence SEQ ID n° 3.

i) TET ou Cro-Oligom (Figure 5D)

20 Ce type de molécule est semblable au type d), à l'exception du domaine de liaison au transactivateur qui est remplacé par le domaine d'oligomérisation à la protéine p53 de séquence SEQ ID n° 3.

j) TET ou Cro-Hinge-Oligom (Figure 5E)

25 Ce type de molécule est semblable au type e), à l'exception du domaine de liaison au transactivateur qui est remplacé par le domaine d'oligomérisation à la protéine p53 de séquence SEQ ID n° 3.

Par ailleurs, dans chacune de ces molécules, le bras peptidique peut être aisément remplacé par la séquence (G4S)3 (SEQ ID n° 5).

30 Un autre objet de la présente invention réside dans une séquence d'acide nucléique codant pour une molécule chimérique telle que définie ci-avant. Il s'agit avantageusement d'une séquence d'ADN, notamment d'ADNc. Il peut également s'agir d'un ARN. Les séquences de l'invention sont généralement construites par assemblage, au sein d'un vecteur de clonage, des séquences codant pour les différents domaines selon les techniques classiques de la biologie moléculaire. Les séquences d'acides nucléiques de l'invention peuvent éventuellement être modifiées par voie chimique,

35



enzymatique ou génétique, en vue de générer des domaines stabilisés, et/ou multifonctionnels, et/ou de taille réduite, et/ou dans le but de favoriser leur localisation dans tel ou tel compartiment intracellulaire. Ainsi, les séquences d'acides nucléiques de l'invention peuvent comprendre des séquences  
 5 codant pour des peptides de localisation nucléaire (NLS). En particulier, il est possible de fusionner les séquences de l'invention avec la séquence codant pour le NLS du virus SV40, dont la séquence peptidique est la suivante : PKKKRKV (SEQ ID n° 9) (Kalderon et al, Cell 39 (1984) 499).

Les séquences nucléiques selon l'invention font avantageusement  
 10 partie d'un vecteur d'expression, qui peut être de nature plasmidique ou virale.

Un autre objet de la présente invention réside dans une protéine de fusion comprenant un domaine transactivateur transcriptionnel et un domaine de liaison spécifique à une molécule donnée, éventuellement reliés par un  
 15 bras peptidique, ainsi que toute séquence d'acide nucléique codant pour une telle fusion. Le domaine transactivateur peut être issu de toute protéine transactivateur transcriptionnel, tel que p53, VP16, EBNA, E6/E7, Tat, etc.

Un autre objet de l'invention consiste en un système conditionnel d'expression de gènes comprenant :

- 20           - une molécule chimérique telle que définie ci-avant et  
              - une cassette d'expression comportant une séquence régulatrice, un promoteur transcriptionnel minimal et ledit gène.

La cassette d'expression contient les éléments nécessaires à l'activation de l'expression du gène par le transactivateur ou complexe  
 25 transactivateur recruté par la molécule bispécifique. Ainsi, la séquence régulatrice est la séquence de liaison à l'ADN de la molécule chimérique exprimée. Lorsque le domaine de liaison à l'ADN de la molécule chimérique est représenté par tout ou partie de TetR, la séquence régulatrice comprend la séquence SEQ ID n° 1 ou un dérivé de celle-ci, éventuellement répétée  
 30 plusieurs fois. Il s'agit préférentiellement de la séquence op2 (comportant 2 motifs Tetop répétés) ou Op7 (comportant 7 motifs Tetop répétés telle que décrite par exemple dans Weinmann et al., The Plant Journal 5 (1994) 559).

De même, lorsque le domaine de liaison à l'ADN de la molécule chimérique est représenté par tout ou partie de Cro, la séquence régulatrice comprend la séquence SEQ ID n° 2 ou un dérivé de celle-ci, éventuellement répétée plusieurs fois. Il s'agit préférentiellement de la séquence OR3. Les dérivés  
 5 des séquences SEQ ID n° 1 et 2 peuvent être toute séquence obtenue par modification de nature génétique (mutation, délétion, addition, répétition, etc) et conservant la capacité de lier spécifiquement une protéine. De tels dérivés ont été décrits dans la littérature (Baumeister et al précitée, Tovar et al., Mol. Gen. Genet. 215 (1988) 76, WO94/04672).

10 Concernant le promoteur transcriptionnel minimal, il s'agit d'un promoteur dont l'activité dépend de la présence d'un transactivateur. De ce fait, en l'absence de la molécule chimérique, le promoteur est inactif et le gène n'est pas ou peu exprimé. En revanche, en présence de la molécule chimérique, le transactivateur ou complexe transactivateur recruté permet  
 15 d'induire l'activité du promoteur minimal et ainsi l'expression du gène d'intérêt. Le promoteur minimal est constitué généralement d'une boîte TATA ou INR. Ces éléments sont en effet les éléments minimum nécessaires à l'expression d'un gène en présence d'un transactivateur. Le promoteur minimal peut être préparé à partir de tout promoteur par modification  
 20 génétique. A titre d'exemple préféré de promoteur candidat on peut citer le promoteur du gène de la thymidine kinase. Des résultats intéressants ont plus précisément été obtenus avec un promoteur minimal dérivant du promoteur TK composé des nucléotides -37 à +19. Le promoteur minimal peut également dériver du CMV humain. En particulier, il peut être constitué  
 25 du fragment compris entre les nucléotides -53 à +75 ou -31 à +75 du CMV. Tout promoteur conventionnel peut toutefois être utilisé tel que par exemple le promoteur des gènes codant pour la chloramphénicol acétyl transférase, la  $\beta$ -galactosidase ou encore la luciférase.

La cassette d'expression est avantageusement constituée des  
 30 éléments suivants :

- Comme séquence régulatrice, une séquence comprenant la séquence SEQ ID n° 1 ou 2 ou un dérivé de celles-ci, éventuellement répétée plusieurs fois,

- Comme promoteur minimal, un promoteur dérivé du promoteur du gène de la thymidine kinase (TK),
- une séquence codante d'intérêt.

5            Encore plus préférentiellement, le promoteur minimal est constitué de la région -37 à +19 du promoteur du gène de la thymidine kinase.

Avantageusement, la cassette d'expression est choisie parmi les cassettes de structure Tetop2.TK-Gène; Tetop7.TK-Gène et OR3.TK-Gène.

10           Un autre aspect de l'invention réside dans un vecteur d'expression comprenant une séquence d'acides nucléiques codant pour une molécule chimérique et une cassette d'expression telles que définies ci-avant. Dans les vecteurs de l'invention, la séquence d'acides nucléiques codant pour la molécule chimérique et la cassette d'expression peuvent être insérées dans la même orientation ou dans les orientations opposées. Par ailleurs, le  
15           vecteur peut être de nature plasmidique ou virale.

Parmi les vecteurs viraux, on peut citer plus préférentiellement les adénovirus, les rétrovirus, les virus de l'herpès ou encore les virus adéno associés. Les virus selon la présente invention sont défectifs, c'est-à-dire incapables de se répliquer de façon autonome dans la cellule cible.  
20           Généralement, le génome des virus défectifs utilisés dans le cadre de la présente invention est donc dépourvu au moins des séquences nécessaires à la réplication dudit virus dans la cellule infectée. Ces régions peuvent être soit éliminées (en tout ou en partie), soit rendues non-fonctionnelles, soit substituées par d'autres séquences et notamment par les séquences de  
25           l'invention. Préférentiellement, le virus défectif conserve néanmoins les séquences de son génome qui sont nécessaires à l'encapsidation des particules virales.

S'agissant plus particulièrement d'adénovirus, différents sérotypes, dont la structure et les propriétés varient quelque peu, ont été  
30           caractérisés. Parmi ces sérotypes, on préfère utiliser dans le cadre de la présente invention les adénovirus humains de type 2 ou 5 (Ad 2 ou Ad 5) ou les adénovirus d'origine animale (voir demande WO94/26914). Parmi les adénovirus d'origine animale utilisables dans le cadre de la présente

invention on peut citer les adénovirus d'origine canine, bovine, murine, (exemple : Mav1, Beard et al., Virology 75 (1990) 81), ovine, porcine, aviaire ou encore simienne (exemple : SAV). De préférence, l'adénovirus d'origine animale est un adénovirus canin, plus préférentiellement un  
 5 adénovirus CAV2 [souche manhattan ou A26/61 (ATCC VR-800) par exemple]. De préférence, on utilise dans le cadre de l'invention des adénovirus d'origine humaine ou canine ou mixte.

Préférentiellement, le génome des adénovirus recombinants de l'invention comprend au moins les ITR et la région d'encapsidation d'un  
 10 adénovirus, et la séquence d'acides nucléiques codant pour une molécule chimérique et une cassette d'expression telles que définies ci-avant. Plus préférentiellement, dans le génome des adénovirus de l'invention, la région E1 au moins est non fonctionnelle. Le gène viral considéré peut être rendu non fonctionnel par toute technique connue de l'homme du  
 15 métier, et notamment par suppression totale, substitution (par exemple par les séquences de l'invention), délétion partielle, ou addition d'une ou plusieurs bases dans le ou les gènes considérés. De telles modifications peuvent être obtenues in vitro (sur de l'ADN isolé) ou in situ, par exemple, au moyens des techniques du génie génétique, ou encore par traitement  
 20 au moyen d'agents mutagènes. D'autres régions peuvent également être modifiées, et notamment la région E3 (WO95/02697), E2 (WO94/28938), E4 (WO94/28152, WO94/12649, WO95/02697) et L5 (WO95/02697). Selon un mode préféré de mise en oeuvre, l'adénovirus selon l'invention comprend une délétion dans les régions E1 et E4. Selon un autre mode de  
 25 réalisation préféré, il comprend une délétion dans région E1 au niveau de laquelle sont insérées la région E4 et les séquences de l'invention (Cf FR94 13355).

Les adénovirus recombinants défectifs selon l'invention peuvent être préparés par toute technique connue de l'homme du métier (Levrero  
 30 et al., Gene 101 (1991) 195, EP 185 573; Graham, EMBO J. 3 (1984) 2917). En particulier, ils peuvent être préparés par recombinaison homologue entre un adénovirus et un plasmide portant entre autre les séquences d'ADN de l'invention (séquence codant pour la molécule chimérique + cassette d'expression). La recombinaison homologue se

produit après co-transfection desdits adénovirus et plasmide dans une lignée cellulaire appropriée. La lignée cellulaire utilisée doit de préférence (i) être transformable par lesdits éléments, et (ii), comporter les séquences capables de compléter la partie du génome de l'adénovirus défectif, de préférence sous forme intégrée pour éviter les risques de recombinaison. A titre d'exemple de lignée, on peut mentionner la lignée de rein embryonnaire humain 293 (Graham et al., J. Gen. Virol. 36 (1977) 59) qui contient notamment, intégrée dans son génome, la partie gauche du génome d'un adénovirus Ad5 (12 %) ou des lignées capables de compléter les fonctions E1 et E4 telles que décrites notamment dans les demandes n° WO 94/26914 et WO95/02697.

Ensuite, les adénovirus qui se sont multipliés sont récupérés et purifiés selon les techniques classiques de biologie moléculaire, comme illustré dans les exemples.

15

Concernant les virus adéno-associés (AAV), il s'agit de virus à ADN de taille relativement réduite, qui s'intègrent dans le génome des cellules qu'ils infectent, de manière stable et site-spécifique. Ils sont capables d'infecter un large spectre de cellules, sans induire d'effet sur la croissance, la morphologie ou la différenciation cellulaires. Par ailleurs, ils ne semblent pas impliqués dans des pathologies chez l'homme. Le génome des AAV a été cloné, séquencé et caractérisé. Il comprend environ 4700 bases, et contient à chaque extrémité une région répétée inversée (ITR) de 145 bases environ, servant d'origine de réplication pour le virus. Le reste du génome est divisé en 2 régions essentielles portant les fonctions d'encapsidation : la partie gauche du génome, qui contient le gène rep impliqué dans la réplication virale et l'expression des gènes viraux; la partie droite du génome, qui contient le gène cap codant pour les protéines de capsid du virus.

30

L'utilisation de vecteurs dérivés des AAV pour le transfert de gènes in vitro et in vivo a été décrite dans la littérature (voir notamment WO 91/18088; WO 93/09239; US 4,797,368, US5,139,941, EP 488 528). Ces demandes décrivent différentes constructions dérivées des AAV, dans lesquelles les gènes rep et/ou cap sont délétés et remplacés par un gène d'intérêt, et leur utilisation pour transférer in vitro (sur cellules en culture)

35

ou in vivo (directement dans un organisme) ledit gène d'intérêt. Les AAV recombinants défectifs selon l'invention peuvent être préparés par co-transfection, dans une lignée cellulaire infectée par un virus auxiliaire humain (par exemple un adénovirus), d'un plasmide contenant les  
5 séquences nucléiques de l'invention (séquence codant pour la molécule chimérique + cassette d'expression) bordée de deux régions répétées inversées (ITR) d'AAV, et d'un plasmide portant les gènes d'encapsidation (gènes rep et cap) d'AAV. Les AAV recombinants produits sont ensuite purifiés par des techniques classiques.

10

Concernant les virus de l'herpès et les rétrovirus, la construction de vecteurs recombinants a été largement décrite dans la littérature : voir notamment Breakfield et al., *New Biologist* 3 (1991) 203; EP 453242, EP178220, Bernstein et al. *Genet. Eng.* 7 (1985) 235; McCormick,  
15 *BioTechnology* 3 (1985) 689, etc.

En particulier, les rétrovirus sont des virus intégratifs, infectant sélectivement les cellules en division. Ils constituent donc des vecteurs d'intérêt pour des applications cancer. Le génome des rétrovirus comprend essentiellement deux LTR, une séquence d'encapsidation et  
20 trois régions codantes (gag, pol et env). Dans les vecteurs recombinants dérivés des rétrovirus, les gènes gag, pol et env sont généralement délétés, en tout ou en partie, et remplacés par une séquence d'acide nucléique hétérologue d'intérêt. Ces vecteurs peuvent être réalisés à partir de différents types de rétrovirus tels que notamment le MoMuLV ("murine  
25 moloney leukemia virus"; encore désigné MoMLV), le MSV ("murine moloney sarcoma virus"), le HaSV ("harvey sarcoma virus"); le SNV ("spleen necrosis virus"); le RSV ("rous sarcoma virus") ou encore le virus de Friend.

Pour construire des rétrovirus recombinants selon l'invention, un  
30 plasmide comportant notamment les LTR, la séquence d'encapsidation et les séquences de l'invention (séquence codant pour la molécule chimérique + cassette d'expression) est généralement construit, puis utilisé pour transfecter une lignée cellulaire dite d'encapsidation, capable d'apporter en trans les fonctions rétrovirales déficientes dans le plasmide.  
35 Généralement, les lignées d'encapsidation sont donc capables d'exprimer

les gènes gag, pol et env. De telles lignées d'encapsidation ont été décrites dans l'art antérieur, et notamment la lignée PA317 (US4,861,719); la lignée PsiCRIP (WO90/02806) et la lignée GP+envAm-12 (WO89/07150). Par ailleurs, les rétrovirus recombinants peuvent  
5 comporter des modifications au niveau des LTR pour supprimer l'activité transcriptionnelle, ainsi que des séquences d'encapsidation étendues, comportant une partie du gène gag (Bender et al., J. Virol. 61 (1987) 1639). Les rétrovirus recombinants produits sont ensuite purifiés par des techniques classiques.

10 Un exemple de construction d'un virus recombinant défectif selon l'invention (rétrovirus) est décrit sur la figure 8. Cette figure souligne un deuxième avantage des constructions selon l'invention qui réside dans l'absence d'expression du gène d'intérêt dans les lignées d'encapsidation. Ces lignées étant dépourvues du transactivateur ou complexe  
15 transactivateur recruté par le système de l'invention, le promoteur est inactif et le gène n'est pas exprimé dans la cellule de production (figure 8A). Ce n'est que lorsque le virus a effectivement infecté une cellule cible, c'est-à-dire une cellule dans laquelle est présent le transactivateur ou complexe transactivateur recruté par le système de l'invention, que le  
20 gène est effectivement exprimé (Figure 8B). Ceci est particulièrement avantageux pour la construction de virus contenant des gènes dont l'expression serait toxique pour les cellules (gènes Grb3-3, IL-2, Toxine diphtérique, etc).

25 Pour la mise en oeuvre de la présente invention, il est tout particulièrement avantageux d'utiliser un adénovirus ou un rétrovirus recombinant défectif. Ces vecteurs possèdent en effet des propriétés particulièrement intéressantes pour le transfert de gènes dans les cellules tumorales.

30 Différents types de vecteurs non-viraux peuvent également être utilisés dans le cadre de l'invention. Le système d'expression conditionnel selon l'invention peut en effet être incorporé dans un agent non viral capable de promouvoir le transfert et l'expression d'acides nucléiques  
35 dans des cellules eucaryotes. Les vecteurs chimiques ou biochimiques

représentent une alternative intéressante aux virus naturels en particulier pour des raisons de commodité, de sécurité et également par l'absence de limite théorique en ce qui concerne la taille de l'ADN à transférer.

5 Ces vecteurs synthétiques ont deux fonctions principales, compacter l'acide nucléique à transférer et promouvoir sa fixation cellulaire ainsi que son passage à travers la membrane plasmique et, le cas échéant, les deux membranes nucléaires. Pour pallier à la nature polyanionique des acides nucléiques, les vecteurs non viraux possèdent tous des charges polycationiques.

10 Parmi les vecteurs synthétiques développés, les polymères cationiques de type polylysine, (LKLK)<sub>n</sub>, (LKKL)<sub>n</sub>, polyéthylène imine et DEAE dextran ou encore les lipides cationiques ou lipofectants sont les plus avantageux. Ils possèdent la propriété de condenser l'ADN et de promouvoir son association avec la membrane cellulaire. Parmi ces  
15 derniers, on peut citer les lipopolyamines (lipofectamine, transfectam, etc) et différents lipides cationiques ou neutres (DOTMA, DOGS, DOPE, etc). Plus récemment, il a été développé le concept de la transfection ciblée, médiée par un récepteur, qui met à profit le principe de condenser l'ADN grâce au polymère cationique tout en dirigeant la fixation du complexe à la  
20 membrane grâce à un couplage chimique entre le polymère cationique et le ligand d'un récepteur membranaire, présent à la surface du type cellulaire que l'on veut greffer. Le ciblage du récepteur à la transferrine, à l'insuline ou du récepteur des asialoglycoprotéines des hépatocytes a ainsi été décrit.

25 La présente invention a également pour objet toute composition pharmaceutique comprenant un vecteur tel que défini ci-avant. Ces compositions peuvent être formulées en vue d'une administration par voie topique, orale, parentérale, intranasale, intraveineuse, intramusculaire, sous-cutanée, intraoculaire, etc. Préférentiellement, la composition selon  
30 l'invention contient des véhicules pharmaceutiquement acceptables pour une formulation injectable. Il peut s'agir en particulier de solutions salines (phosphate monosodique, disodique, chlorure de sodium, potassium, calcium ou magnésium, etc, ou des mélanges de tels sels), stériles, isotoniques, ou de compositions sèches, notamment lyophilisées, qui, par addition selon le  
35 cas d'eau stérilisée ou de sérum physiologique, permettent la constitution de



solutés injectables. S'agissant des rétrovirus, ils peuvent être avantageux d'utiliser directement les cellules d'encapsulation ou des cellules infectées ex vivo en vue de leur réimplantation in vivo, éventuellement sous forme de néo-organes (WO94/24298).

5 Les doses de vecteur utilisées pour l'injection peuvent être adaptées en fonction de différents paramètres, et notamment en fonction du mode d'administration utilisé, de la pathologie concernée ou encore de la durée du traitement recherchée. D'une manière générale, les virus recombinaux selon l'invention sont formulés et administrés sous forme de  
10 doses comprises entre  $10^4$  et  $10^{14}$  pfu/ml. Pour les AAV et les adénovirus, des doses de  $10^6$  à  $10^{10}$  pfu/ml peuvent également être utilisées. Le terme pfu ("plaque forming unit") correspond au pouvoir infectieux d'une suspension de virions, et est déterminé par infection d'une culture cellulaire appropriée, et mesure, généralement après 48 heures,  
15 du nombre de plaques de cellules infectées. Les techniques de détermination du titre pfu d'une solution virale sont bien documentées dans la littérature.

Le système d'expression selon l'invention et les vecteurs  
20 correspondants sont particulièrement utiles pour contrôler l'expression de gènes d'intérêt dans le cadre de thérapies cellulaires ou géniques. Ils peuvent ainsi être utilisés pour contrôler l'expression de toute séquence codante d'intérêt, et notamment d'une séquence codant pour un produit thérapeutique, qu'il s'agisse d'un peptide, polypeptide, protéine, acide ribonucléiques, etc. Plus particulièrement, le gène est une séquence  
25 d'ADN (ADNc, ADNg, ADN synthétique, humain, animal, végétal, etc) codant pour un produit protéique tel que des enzymes, les dérivés sanguins, les hormones, les lymphokines : interleukines, interférons, TNF, etc (FR 9203120), les facteurs de croissance, les neurotransmetteurs ou  
30 leurs précurseurs ou enzymes de synthèse, les facteurs trophiques : BDNF, CNTF, NGF, IGF, GMF, aFGF, bFGF, NT3, NT5, etc; les apolipoprotéines : ApoA1, ApoAIV, ApoE, etc (FR 93 05125), la dystrophine ou une minidystrophine (FR 9111947), les gènes suppresseurs de tumeurs : p53, Rb, Rap1A, DCC, k-rev, etc (FR 93  
35 04745), les gènes codant pour des facteurs impliqués dans la coagulation

: Facteurs VII, VIII, IX, etc, ou encore tout ou partie d'une immunoglobuline naturelle ou artificielle (Fab, ScFv, etc), un ARN ligand (WO91/19813) etc.

Le gène d'intérêt peut également être une séquence antisens, dont l'expression dans la cellule cible permet de contrôler l'expression de gènes ou la transcription d'ARNm cellulaires. De telles séquences peuvent par exemple être transcrites, dans la cellule cible, en ARN complémentaires d'ARNm cellulaires et bloquer ainsi leur traduction en protéine, selon la technique décrite dans le brevet EP 140 308.

La présente invention est particulièrement adaptée à l'expression de séquences codant pour des facteurs toxiques. Il peut s'agir en particulier de poisons pour les cellules (toxine diphtérique, toxine pseudomonas, ricine A, etc) de produit induisant une sensibilité à un agent externe (gènes suicides :Thymidine kinase, cytosine désaminase, etc) ou encore de gènes tueurs capables d'induire la mort cellulaire (Grb3-3 (PCT/FR94/00542), ScFv anti-ras (WO94/29446), etc). Le système de l'invention permet en effet de produire des vecteurs notamment viraux contenant ces séquences sans toxicité pour les cellules de production, puis ensuite d'induire l'expression de ces molécules toxiques sélectivement dans des cellules cibles présentant le transactivateur ou complexe transactivateur désiré. Ce type de construction est donc particulièrement adapté à des stratégies de thérapies antitumorales par exemple, dans lesquelles l'objectif est de détruire sélectivement les cellules affectées. Ce système est également particulièrement intéressant pour l'expression de cytokines, interférons, TNF ou TGF par exemple, dont une production incontrôlée peut avoir des effets secondaires très marqués.

La présente demande sera décrite plus en détail à l'aide des exemples qui suivent, qui doivent être considérés comme illustratifs et non limitatifs.

### 30 Légende des Figures

Figure 1: Représentation du système d'expression conditionnel selon l'invention permettant le recrutement sélectif d'un transactivateur au moyen d'un domaine d'oligomérisation (A) ou d'un ScFv (B).

Figure 2: Représentation du système d'expression conditionnel selon l'invention permettant le recrutement sélectif d'un complexe transactivateur.

Figure 3: Représentation d'une cassette d'expression selon l'invention comportant une séquence régulatrice Tetop7, un promoteur minimal (boîte TATA) et un gène (CAT).

Figure 4: Représentation d'une cassette d'expression selon l'invention comportant une séquence régulatrice OR3, un promoteur minimal (boîte TATA) et un gène (CAT).

Figure 5: Représentation de molécules chimériques bispécifiques selon l'invention.

Figure 6: Construction de séquences d'ADN codant pour des molécules chimériques bispécifiques selon l'invention.

Figure 7: Représentation de constructions chimériques contrôle.

Figure 8: Structure et fonctionnement d'un vecteur viral (rétrovirus) selon l'invention.

### **Techniques générales de Biologie Moléculaire**

Les méthodes classiques de biologie moléculaires telles que la centrifugation d'ADN plasmidique en gradient de chlorure de césium-bromure d'éthidium, les digestions par les enzymes de restriction, l'électrophorèse sur gel, la transformation dans *E.coli*, la précipitation des acides nucléiques etc, sont décrites dans la littérature (Maniatis *et al.*, 1989).

Les enzymes ont été fournies par New-England Biolabs (Beverly, MA). Pour les ligatures les fragments d'ADN sont séparés selon leur taille sur des gels d'agarose de 0,8 à 1,5 %, purifiés par GeneClean (BIO101, LaJolla CA) et incubés de nuit à 14°C dans un tampon Tris-HCl pH 7.4 50 mM, MgCl<sub>2</sub> 10 mM, DTT 10 mM, ATP 2 mM, en présence d'ADN ligase du phage T4.

L'amplification par PCR, (Polymerase Chain Reaction), a également été réalisée selon Maniatis *et al.*, 1989, avec les spécifications suivantes :

- Concentration en MgCl<sub>2</sub> portée à 8mM.
- Température de dénaturation 95°C, température d'hybridation 55°C, température d'allongement 72°C. Ce cycle a été répété 25 fois dans un PE9600 Thermalcycler (Perkin Elmer, Norwalk CO).

Les oligonucléotides sont synthétisés en utilisant la chimie des phosphoramidites protégés en β par un groupement cyanoéthyl, (Sinha *et*

al.,1984, Giles 1985), avec le synthétiseur automatique d'ADN Applied Biosystem modèle 394, (Applied Biosystem, Foster City CA), selon les recommandations du fabricant.

5 Le séquençage a été effectué sur des matrices double-brin par la méthode de terminaison de chaînes en utilisant des amorces fluorescentes. Nous avons utilisé le kit de séquençage Taq Dye Primer Kit de chez Applied Biosystem (Applied Biosystem, Foster City CA) selon les spécifications du fabricant.

### Exemples

10

Exemple 1 : Construction de cassettes d'expression comportant une séquence de régulation, un promoteur transcriptionnel minimal et un gène.

#### 1.1. Construction du plasmide pTETop7/CAT

15

Le plasmide pTETop7/CAT contient les éléments suivants (Figure 3) :

- Une séquence de régulation constituée d'une séquence d'interaction avec le répresseur tétracycline tetR composée de 7 motifs Tetop (SEQ ID n° 1) répétés;

20

- un promoteur minimal dérivé du promoteur du gène de la thymidine kinase (région -37 à + 19 portant la boîte TATA);

- la séquence codant pour la chloramphénicol acétyl transférase (CAT) sous contrôle dudit promoteur minimal.

25 Ce plasmide a été construit par clonage du fragment SmaI-BglII du plasmide pUHD10-7 (WO94/29442) dans le plasmide pKK232-8 (Pharmacia) préalablement digéré par SmaI et BamHI.

#### 1.2. Construction du plasmide pOR3/CAT

30

Le plasmide pOR3/CAT contient les éléments suivants (Figure 4) :

- Une séquence de régulation constituée d'une séquence OR3 d'interaction avec le répresseur Cro (SEQ ID n° 2);

- un promoteur minimal dérivé du promoteur du gène de la thymidine kinase (région -37 à + 19 portant la boîte TATA);

- la séquence codant pour la chloramphénicol acétyl transférase (CAT) sous contrôle dudit promoteur minimal.

Ce plasmide a été construit de la manière suivante : La séquence OR3 d'interaction avec le répresseur Cro a été synthétisée artificiellement. Pour  
5 cela, les deux oligonucléotides suivants ont été synthétisés :

Oligo 5533 (SEQ ID n° 22) : 5'-GATCCTATCACCGCAAGGGATAA-3'

Oligo 5534 (SEQ ID n° 23) : 3'-GATAGTGGCGTTCCTATTTCTGA-5'

Ces deux oligonucléotides ont ensuite été hybridés pour reconstituer  
10 la séquence double brin OR3 bordée de séquences permettant son clonage orienté comme suit :

GATCCTATCACCGCAAGGGATAA

GATAGTGGCGTTCCTATTTCTGA

15

### 1.3. Construction de cassettes d'expression de gènes toxiques

Les cassettes d'expression de gènes toxiques sont obtenues à partir des plasmides décrits ci-dessus (1.1. et 1.2.) par remplacement de la  
20 séquence CAT par la séquence codant pour le produit toxique, de préférence le gène Grb3-3 (PCT/FR94/00542), le gène de la thymidine kinase, le gène codant pour la toxine diphtérique ou pseudomonas, etc.

#### Exemple 2 : Construction d'un anticorps simple chaîne spécifique de p53

25

Cet anticorps simple chaîne a été construit selon le protocole suivant :

- Les cDNA codant pour les régions VH et VL ont été obtenus à partir de l'hybridome pAb421 produisant un anticorps anti-p53. Pour cela, les ARN totaux de l'hybridome ont été extraits et soumis à une réaction de  
30 transcription inverse en utilisant des hexamères random comme amorces. L'utilisation de ce type d'amorce permet d'éviter l'emploi d'amorces spécifiques des immunoglobulines. Les clones d'ADNc obtenus ont une longueur suffisante pour cloner les régions V. Toutefois, dans la mesure où ils représentent une faible fraction des cDNA totaux présents, une réaction  
35 préalable d'amplification doit être réalisée pour produire suffisamment de

DNA pour le clonage. Pour cela, les ADNc codant pour les régions VH et VL ont été amplifiés séparément. Les amorces utilisées sont des oligonucléotides hybridant au niveau des extrémités opposées des régions variables de chaque chaîne (H et L). Le produit d'amplification utilisant les amorces spécifiques des chaînes lourdes est un fragment de 340 paires de bases environ. Le produit d'amplification utilisant les amorces spécifiques des chaînes légères est un fragment de 325 paires de bases environ.

- Après purification, les ADNc codant pour les régions VH et VL de l'anticorps ont été assemblés en une chaîne unique au moyen d'un bras nucléotidique (L). Le bras nucléotidique a été construit de telle sorte que l'une des extrémités se lie à l'extrémité 3' de l'ADNc codant pour la région VH et l'autre à l'extrémité 5' de l'ADNc codant pour la région VL. La séquence du bras code pour le peptide SEQ ID n° 5. La séquence assemblée de 700 pb environ contient, sous forme d'un fragment NcoI-NotI, l'enchaînement VH-L-VL dont la séquence est représentée SEQ ID n° 4 (acides aminés 9 à 241). Cette séquence inclut également en C-terminal la séquence tag de myc (résidus 256 à 266).

### Exemple 3 : Construction de séquences d'acides nucléiques codant pour des molécules chimériques bispécifiques contenant un domaine de liaison à un transactivateur constitué d'un anticorps simple chaîne (ScFv).

#### 3.1. Construction d'un plasmide comportant une séquence ScFv-myc-Hinge-TetR ou Cro (Figures 5A et 6)

Le fragment NcoI-NotI contenant l'ADNc codant pour le ScFv anti-p53 a tout d'abord été cloné dans un plasmide de type pUC19. La séquence codant pour l'épitope VSV (SEQ ID n° 7) ou myc (SEQ ID n° 8) est insérée en aval du fragment (Figure 6).

Les séquences codant pour les protéines TetR et Cro ont ensuite été obtenues comme suit :

- La séquence codant pour TetR a été obtenue par amplification à partir d'un plasmide matrice portant la séquence tetR au moyen des oligonucléotides suivants :

Oligo 5474 (SEQ ID n° 10) :

GGCTCTAGACCCAAGCCCAGTACCCCCCAGGTTCTTCAACGCG  
TGGATCCATGTCCAGATTAGATAAAAGTAAAG

Oligo 5475 (SEQ ID n° 11) :

CGTACGGAATTCGGGGCCCTTACTCGAGGGACCCACTTTTACATTT  
5 AAGTTG

Ces oligonucléotides apportent également la séquence codant pour le bras peptidique Hinge reliant les deux domaines fonctionnels de la molécule.

Le fragment amplifié contient donc la séquence codant pour le bras  
10 peptidique et pour le domaine de liaison à l'ADN tetR. Ce fragment a été cloné aux sites XbaI-EcoRI du plasmide obtenu ci-dessus pour générer un plasmide contenant la séquence codant pour la molécule ScFv-myc-Hinge-TetR (Figure 6).

- La séquence codant pour Cro a été obtenue par amplification sur  
15 une matrice d'ADN du bactériophage Lambda au moyen des oligonucléotides suivants :

Oligo 5531 (SEQ ID n° 12) :

GGCTCTAGACCCAAGCCCAGTACCCCCCAGGTTCTTCAACGCG  
TGGATCCATGGAACAACGCATAACCCTGAAAG

20 Oligo 5532 (SEQ ID n° 13) :

CGTACGGAATTCGGGGCCCTTACTCGAGTGCTGTTGTTTTTTTGT  
ACTCGG

Ces oligonucléotides apportent également la séquence codant pour le bras peptidique Hinge reliant les deux domaines fonctionnels de la  
25 molécule.

Le fragment amplifié contient donc la séquence codant pour le bras peptidique et pour le domaine de liaison à l'ADN Cro. Ce fragment a été cloné aux sites XbaI-EcoRI du plasmide obtenu ci-dessus pour générer un plasmide contenant la séquence codant pour la molécule ScFv-myc-Hinge-Cro  
30 (Figure 6).

3.2. Construction d'un plasmide comportant une séquence ScFv-Hinge-TetR ou Cro (Figure 5B)

Cet exemple décrit la construction de plasmides portant une séquence codant pour une molécule chimérique bispécifique selon l'invention dépourvue de séquence tag.

5 Ces plasmides ont été obtenus à partir des plasmides décrits en 3.1. ci-avant par digestion par les enzymes NotI et XbaI. Cette digestion permet d'exciser le fragment portant la région codant pour le tag myc.

3.3. Construction d'un plasmide comportant une séquence ScFv-TetR ou Cro (Figure 5C)

10 Cet exemple décrit la construction de plasmides portant une séquence codant pour une molécule chimérique bispécifique selon l'invention dépourvue de bras et de séquence tag.

15 Ces plasmides ont été obtenus à partir des plasmides décrits en 3.1. ci-avant par digestion par les enzymes NotI et BamHI. Cette digestion permet d'exciser le fragment portant la région codant pour le tag myc et pour le bras peptidique Hinge.

20 Exemple 4 : Construction de séquences d'acides nucléiques codant pour des molécules chimériques bispécifiques contenant un domaine de liaison à un transactivateur constitué d'un domaine d'oligomérisation.

4.1. Clonage de la région d'oligomérisation de la protéine p53 (fragment 320-393).

25 L'ADNc codant pour la région d'oligomérisation de la protéine p53 (SEQ ID n° 3) a été obtenu par amplification par PCR sur un plasmide portant le cDNA de la p53 sauvage humaine à l'aide des oligonucléotides suivants :

Oligo 5535 (SEQ ID n° 14) :

CAGGCCATGGCATGAAGAAACCACTGGATGGAGAA

30 (la partie soulignée représente un site NcoI)

Oligo 5536 (SEQ ID n° 15) :

CGTCGGATCCTCTAGATGCGGCCGCGTCTGAGTCAGGCCCTTC

(Partie soulignée : site BamHI; Double-souligné : site XbaI; Gras : site NotI).

35



4.2. Les plasmides p53 320/393-myc-Hinge-TetR ou Cro (Figure 5A) ont été obtenus par clonage du fragment amplifié ci-dessus sous forme d'un fragment NcoI-NotI dans les sites correspondants des plasmides décrits dans l'exemple 3.1., en substitution de la région codant pour le ScFv.

5

4.3. Les plasmides p53 320/393-Hinge-TetR ou Cro (Figure 5B) ont été obtenus par clonage du fragment amplifié en 4.1. sous forme d'un fragment NcoI-XbaI dans les sites correspondants des plasmides décrits dans l'exemple 3.1., en substitution de la région codant pour le ScFv et le tag.

10

4.4. Les plasmides p53 320/393-TetR ou Cro (Figure 5C) ont été obtenus par clonage du fragment amplifié en 4.1. sous forme d'un fragment NcoI-BamHI dans les sites correspondants des plasmides décrits dans l'exemple 3.1., en substitution de la région codant pour le ScFv, le tag et le Hinge.

15

4.5. Les plasmides tetR ou Cro- p53 320/393 (Figure 5D) ou tetR ou Cro-Hinge-p53 320/393 (Figure 5E) ont été obtenus par clonage de fragments amplifiés par PCR sur un plasmide portant le cDNA de la p53 humaine sauvage à l'aide des oligos 5537/5539 ou 5538/5539 digérés par XhoI/EcoRI dans les plasmides décrits en 3.1., préalablement digérés par XhoI/EcoRI.

20

Oligo 5537 (SEQ ID n° 16) :

25

CAGGCTCGAGAAGAAACCACTGGATGGAGAA

Oligo 5538 (SEQ ID n° 17) :

CAGGCTCGAGCCCAAGCCCAGTACCCCCCAGGTTCTTCAAAGA

AACCACTGGATGGAGAA

Oligo 5539 (SEQ ID n° 18) :

30

GGTCGAATTCGGGCCCTCAGTCTGAGTCAGGCCCTTC

Exemple 5 : Construction d'un plasmide contrôle portant une séquence codant pour une molécule chimérique comportant un domaine de liaison à l'ADN (TetR ou Cro) et le domaine transactivateur de la protéine p53 (région 1-73).

35

Les plasmides p53 1/73 - TetR ou Cro avec ou sans tag (myc ou VSV) et Hinge (Figures 7 A, B et C) ont été obtenus par clonage de fragments amplifiés par PCR à partir d'un plasmide portant le cDNA de la p53 sauvage humaine à l'aide des oligos 5661/5662 puis digérés par NcoI/NotI, NcoI/XbaI, NcoI/BamHI dans les plasmides décrits en 3.1., préalablement digérés par NcoI/NotI, NcoI/XbaI ou NcoI/BamHI.

Oligo 5661 (SEQ ID n° 19) :

CAGGCCATGGAGGAGCCGCAGTCAGATCC

Oligo 5662 (SEQ ID n° 20) :

CGTCGGATCCTCTAGATGCGGCCGCCACGGGGGGAGCAGCCTC

TGG

## LISTE DE SEQUENCES

SEQ ID n° 1 :

TCTCTATCACTGATAGGGA

5 SEQ ID n° 2 :

TATCACCGCAAGGGATA

SEQ ID n° 3 :

KKPLDGEYFTLQIRGRERFEMFRELNEALELKDAQAGKEPGGSRAHSSHLKSKKGQSTSRH

KKLMFKTEGPDS

10 SEQ ID n° 4 :

1/1 *PELIS LEADER SEIT NOE* *V<sub>H</sub> →* 31/11 *Pat I*  
 CTC GCG GCC CAG CCG GCC ATG GCC [CAG GTG AAA CTG CAG GAG TCA GGG GCA GAG CTT GTG  
 leu ala ala gln pro ala met ala gln val lys leu gln glu ser gly ala glu leu val  
 61/21 91/31 *CDR1*  
 GAG TCA GGG GCC TCA GTC AAG TTG TCC TGC ACA GCT TCT GGC TTC AAC ATT AAA GAC TAC  
 15 glu ser gly ala ser val lys leu ser cys thr ala ser gly phe asn ile lys asp tyr  
 121/41 151/51  
 TAT ATG CAG TGG GTG AAG CAG AGG CCT GAA CAG GGC CTG GAG TGG ATT GGA TGG ATT GAT  
tyr met his trp val lys gln arg pro glu gln gly leu glu trp ile gly trp ile asp  
 181/61 *CDR2* 211/71  
 CCT GAG AAT GGT GAT ACT GAA TAT GCC CCG AAG TTC CAG GGC AAG GCC ACT ATG ACT GCA  
 pro glu asn gly asp thr glu tyr ala pro lys phe gln gly lys ala thr met thr ala  
 241/81 271/91  
 20 GAC ACA TCC TCC AAT ACA GCC TAC CTG CAG CTC AGC AGC CTG GCA TCT GAG GAC ACT GCC  
 asp thr ser ser asn thr ala tyr leu gln leu ser ser leu ala ser glu asp thr ala  
 301/101 *CDR3* 331/111 *Set II*  
 GTC TAT TAT TGT AAT TTT TAC GGG GAT GCT TTG GAC TAC TGC GGC CAA GGG ACC ACG GTC  
 val tyr tyr cys asn phe tyr gly asp ala leu asp tyr trp gly gln gly thr thr val *V<sub>K</sub> →*  
 361/121 391/131 *LINKER*  
 ACC GTC TCC TCA GGT GGA GGC GGT TCA GGC GGA GGT GGC TCT GGC GGT GGC GGA TCG GAT  
 25 thr val ser ser gly gly gly gly ser gly gly gly ser gly gly gly ser asp  
 421/141 451/151  
 GTT TTG ATG ACC CAA ACT CCA CTC ACT TTG TCG GTT ACC ATT GGA CAA CCA GCC TCC ATC  
 val leu met thr gln thr pro leu thr leu ser val thr ile gly gln pro ala ser ile  
 481/161 511/171 *CDR1*  
 TCT TGC AAG TCA AGT CAG AGC CTC TTG GAT AGT GAT GGA AAG ACA TAT TTG AAT TGG TTG  
 ser cys lys ser ser gln ser leu leu asp ser asp gly lys thr tyr leu asn trp leu  
 541/181 571/191 *CDR2*  
 30 TTA CAG AGG CCA GGC CAG TCT CCA AAG CCG CTA ATC TAT CTG GTG TCT AAA CTG GAC TCT  
 leu gln arg pro gly gln ser pro lys arg leu ile tyr leu val ser lys leu asp ser  
 601/201 631/211  
 GGA GTC CCT GAC AGG TTC ACT GGC AGT GGA TCA GGG ACA GAT TTC ACA CTG AAA ATC AAC  
 gly val pro asp arg phe thr gly ser gly ser gly thr asp phe thr leu lys ile asn  
 661/221 691/231 *CDR3*  
 AGA GTG GAG GCT GAG GAT TTG GGA GTT TAT TAT TGC TGG CAA GGT ACA CAT TCT CCG CTC  
 arg val glu ala glu asp leu gly val tyr tyr cys trp gln gly thr his ser pro leu  
 721/241 751/251 *Not 1* *NYC-TAG*  
 35 ACC TTC GGT GCT GGG ACC AAG CTG GAG CTG AAA CGG GCC GCA GAA CAA AAA CTC ATC  
thr phe gly ala gly thr lys leu glu leu lys arg ala ala glu gln lys leu ile  
 781/261 811/271  
 TCA GAA GAG GAT CTG AAT TAA TAA GAA TTC ACT GGC  
 ser glu glu asp leu asn OCH OCH glu phe thr gly

SEQ ID n° 5 :  
GGGSGGGSGGGGS

SEQ ID n° 6 :  
5 CCC AAG CCC AGT ACC CCC CCA GGT TCT TCA  
P K P S T P P G S S

SEQ ID n° 7 :  
ATG.AAC CGG CTG GGC AAG  
M N R L G K

SEQ ID n° 8 :  
10 GAA CAA AAA CTC ATC TCA GAA GAG GAT CTG AAT  
E Q K L I S E E D L N

SEQ ID n° 9 :  
PKKKRKV

SEQ ID n° 10 :  
15 GGCTCTAGACCCAAGCCCAGTACCCCCCAGGTTCTTCAACGCGTGGATCCATGTCCAGAT  
TAGATAAAAAGTAAAG

SEQ ID n° 11 :  
CGTACGGAATTTCGGGCCCTTACTCGAGGGACCCACTTTCACATTTAAGTTG

SEQ ID n° 12 :  
20 GGCTCTAGACCCAAGCCCAGTACCCCCCAGGTTCTTCAACGCGTGGATCCATGGAACAAC  
GCATAACCCTGAAAG

SEQ ID n° 13 :  
CGTACGGAATTTCGGGCCCTTACTCGAGTGCTGTTGTTTTTTTGTACTCGG

SEQ ID n° 14 :  
25 CAGGCCATGGCATGAAGAAACCACTGGATGGAGAA

SEQ ID n° 15 :  
CGTCGGATCCTCTAGATGCGGCCGCGTCTGAGTCAGGCCCTTC

SEQ ID n° 16 :  
CAGGCTCGAGAAGAAACCACTGGATGGAGAA

SEQ ID n° 17 :  
30 CAGGCTCGAGCCCAAGCCCAGTACCCCCCAGGTTCTTCAAAGAAACCACTGGATGGAGAA

SEQ ID n° 18 :  
GGTCGAATTTCGGGCCCTCAGTCTGAGTCAGGCCCTTC

SEQ ID n° 19 :  
35 CAGGCCATGGAGGAGCCGCAGTCAGATCC

SEQ ID n° 20 :

CGTCGGATCCTCTAGATGCGGCCGCCACGGGGGAGCAGCCTCTGG

SEQ ID n° 21 :

Met Glu Gln Arg Ile Thr Leu Lys Asp Tyr Ala Met Arg Phe Gly  
Gln Thr Lys Thr Ala Lys Asp Leu Gly Val Tyr Gln Ser Ala Ile  
5 Asn Lys Ala Ile His Ala Gly Arg Lys Ile Phe Leu Thr Ile Asn  
Ala Asp Gly Ser Val Tyr Ala Glu Glu Val Lys Pro Phe Pro Ser  
Asn Lys Lys Thr Thr Ala

SEQ ID n° 22 :

GATCCTATCACCGCAAGGGATAA

10 SEQ ID n° 23 :

3'-GATAGTGGCGTTCCTATTTCGA-5'

## REVENDEICATIONS

1. Molécule chimérique bispécifique comprenant un domaine capable de lier sélectivement une séquence définie d'ADN et un domaine détecteur capable de lier spécifiquement un transactivateur ou transrépresseur ou un complexe transactivateur ou transrépresseur.  
5
2. Molécule selon la revendication 1 caractérisée en ce que le domaine capable de lier sélectivement une séquence définie d'ADN dérive d'une protéine capable d'interagir avec l'ADN.
3. Molécule selon la revendication 2 caractérisée en ce que le domaine capable de lier sélectivement une séquence définie d'ADN dérive d'une protéine eucaryote.  
10
4. Molécule selon la revendication 3 caractérisée en ce que le domaine capable de lier sélectivement une séquence définie d'ADN dérive des protéines p53, STAT ou NFkB.
5. Molécule selon la revendication 2 caractérisée en ce que le domaine capable de lier sélectivement une séquence définie d'ADN dérive d'une protéine procaryote.  
15
6. Molécule selon la revendication 5 caractérisée en ce que la protéine procaryote est un répresseur bactérien.
7. Molécule selon la revendication 6 caractérisée en ce que le domaine capable de lier sélectivement une séquence définie d'ADN dérive la protéine tetR.  
20
8. Molécule selon la revendication 6 caractérisée en ce que le domaine capable de lier sélectivement une séquence définie d'ADN dérive la protéine Cro.  
25
9. Molécule selon l'une des revendications 2 à 8 caractérisée en ce que le domaine capable de lier sélectivement une séquence définie d'ADN comprend le domaine d'interaction avec l'ADN de ladite protéine.

10. Molécule selon l'une des revendications 2 à 8 caractérisée en ce que le domaine capable de lier sélectivement une séquence définie d'ADN est constitué de la protéine complète.
- 5 11. Molécule selon la revendication 10 caractérisée en ce que le domaine capable de lier sélectivement une séquence définie d'ADN est constitué de la protéine tetR.
12. Molécule selon la revendication 10 caractérisée en ce que le domaine capable de lier sélectivement une séquence définie d'ADN est constitué de la protéine Cro.
- 10 13. Molécule selon la revendication 1 caractérisée en ce que le domaine capable de lier spécifiquement le transactivateur ou transrépresseur ou le complexe transactivateur ou transrépresseur est un domaine d'oligomérisation.
- 15 14. Molécule selon la revendication 13 caractérisée en ce que le domaine d'oligomérisation est un leucine-zipper, un domaine SH3 ou SH2.
15. Molécule selon la revendication 13 caractérisée en ce que le domaine d'oligomérisation capable de lier spécifiquement le transactivateur est constitué de la partie C-terminale de la protéine p53.
- 20 16. Molécule selon la revendication 15 caractérisée en ce que le domaine d'oligomérisation est constitué de la partie C-terminale de la protéine p53 allant du résidu 320 à 393 (SEQ ID n° 3), 302-360 ou 302-390.
- 25 17. Molécule selon la revendication 1 caractérisée en ce que le domaine capable de lier spécifiquement le transactivateur ou transrépresseur ou le complexe transactivateur ou transrépresseur est un domaine synthétique ou naturel connu pour interagir avec ledit transactivateur ou transrépresseur ou complexe transactivateur ou transrépresseur.
18. Molécule selon la revendication 1 caractérisée en ce que le domaine capable de lier spécifiquement le transactivateur ou transrépresseur ou le complexe transactivateur ou transrépresseur est un anticorps ou un fragment

ou dérivé d'anticorps dirigé contre le transactivateur ou transrépresseur ou complexe transactivateur ou transrépresseur.

19. Molécule selon la revendication 18 caractérisée en ce que le domaine capable de lier spécifiquement le transactivateur ou le complexe transactivateur est constitué par un fragment Fab ou F(ab)'<sub>2</sub> d'anticorps ou une région VH ou VL d'un anticorps.
20. Molécule selon la revendication 18 caractérisée en ce que le domaine capable de lier spécifiquement le transactivateur ou le complexe transactivateur est constitué par un anticorps simple chaîne (ScFv) comprenant une région VH liée à une région VL par un bras.
21. Molécule selon la revendication 1 caractérisée en ce que le domaine de liaison à l'ADN et le domaine de liaison au transactivateur sont reliés entre eux par l'intermédiaire d'un bras.
22. Molécule selon la revendication 21 caractérisée en ce que le bras est constitué d'un peptide comprenant de 5 à 30 acides aminés et, de préférence, de 5 à 20 acides aminés.
23. Molécule selon la revendication 22 caractérisée en ce que le bras est choisi parmi les peptides de séquence SEQ ID n° 5 ou SEQ ID n° 6.
24. Molécule selon l'une des revendications précédentes caractérisée en ce que le domaine de liaison à l'ADN est situé en position N-terminale et le domaine de liaison au transactivateur est situé en position C-terminale.
25. Molécule selon l'une des revendications 1 à 23 caractérisée en ce que le domaine de liaison à l'ADN est situé en position C-terminale et le domaine de liaison au transactivateur est situé en position N-terminale.
26. Molécule chimérique bispécifique de structure ScFv-VSV/myc-Hinge-TET ou Cro (Figure 5A).
27. Molécule chimérique bispécifique de structure ScFv-Hinge-TET ou Cro (Figure 5B).



28. Molécule chimérique bispécifique de structure ScFv-TET ou Cro (Figure 5C).
29. Molécule chimérique bispécifique de structure TET ou Cro-ScFv (Figure 5D).
30. Molécule chimérique bispécifique de structure TET ou Cro-Hinge-ScFv (Figure 5E).
31. Molécule chimérique bispécifique de structure Oligom-VSV/myc-Hinge-TET ou Cro (Figure 5A), Oligom-Hinge-TET ou Cro (Figure 5B) ou Oligom-TET ou Cro (Figure 5C).
32. Séquence d'acide nucléique codant pour une molécule chimérique selon l'une des revendications 1 à 31.
33. Séquence d'acide nucléique selon la revendication 32 caractérisée en ce qu'il s'agit d'une séquence d'ADN.
34. Séquence d'acide nucléique selon la revendication 32 ou 33 caractérisée en ce qu'elle fait partie d'un vecteur.
35. Système conditionnel d'expression de gènes comprenant :
- une molécule chimérique telle que définie dans les revendications 1 à 31, et,
  - une cassette d'expression comportant une séquence régulatrice, un promoteur transcriptionnel minimal et ledit gène.
36. Système conditionnel selon la revendication 35 caractérisé en ce que le domaine de liaison à l'ADN de la molécule chimérique est représenté par tout ou partie de TetR et la séquence régulatrice comprend la séquence SEQ ID n° 1 ou un dérivé de celle-ci, éventuellement répétée plusieurs fois.
37. Système conditionnel selon la revendication 35 caractérisé en ce que le domaine de liaison à l'ADN de la molécule chimérique est représenté par tout ou partie de Cro et la séquence régulatrice comprend la séquence SEQ ID n° 2 ou un dérivé de celle-ci, éventuellement répétée plusieurs fois.

38. Système conditionnel selon l'une des revendications 35 à 37 caractérisé en ce que le promoteur minimal comprend une boîte TATA ou INR.
39. Système conditionnel selon la revendication 38 caractérisé en ce que le promoteur minimal dérive du promoteur du gène de la thymidine kinase.
- 5 40. Système conditionnel selon la revendication 39 caractérisé en ce que le promoteur minimal est composé des nucléotides -37 à +19 du promoteur du gène de la thymidine kinase.
41. Vecteur comprenant :
- 10       - une séquence d'acide nucléique codant pour une molécule chimérique selon l'une des revendications 1 à 31, et
- une cassette d'expression comportant une séquence de régulatrice, un promoteur transcriptionnel minimal et une séquence codante d'intérêt.
- 42- Vecteur selon la revendication 41 caractérisé en ce que le promoteur transcriptionnel minimal est défini selon les revendications 38 à 40.
- 15 43. Vecteur selon la revendication 41 caractérisé en ce que le domaine de liaison à l'ADN de la molécule chimérique est représenté par tout ou partie de TetR et la séquence régulatrice comprend la séquence SEQ ID n° 1 ou un dérivé de celle-ci, éventuellement répétée plusieurs fois.
- 20 44. Vecteur selon la revendication 41 caractérisé en ce que le domaine de liaison à l'ADN de la molécule chimérique est représenté par tout ou partie de Cro et la séquence régulatrice comprend la séquence SEQ ID n° 2 ou un dérivé de celle-ci, éventuellement répétée plusieurs fois.
45. Vecteur selon l'une des revendications 41 à 44 caractérisé en ce que la
- 25 séquence codante d'intérêt est une séquence d'ADN codant pour un produit thérapeutique.
46. Vecteur selon la revendication 45 caractérisé en ce que le produit thérapeutique est un peptide ou polypeptide toxique.
47. Vecteur selon la revendication 46 caractérisé en ce que le produit
- 30 thérapeutique toxique est choisi parmi la toxine diphtérique, la toxine

pseudomonas, la ricine A, la thymidine kinase, la cytosine désaminase, la protéine Grb3-3, ou le ScFv Y28.

48. Vecteur selon l'une des revendications 41 à 47 caractérisé en ce qu'il s'agit d'un vecteur plasmidique.

5 49. Vecteur selon l'une des revendications 41 à 47 caractérisé en ce qu'il s'agit d'un vecteur viral.

50. Vecteur selon la revendication 49 caractérisé en ce qu'il s'agit d'un adénovirus recombinant défectif.

10 51. Vecteur selon la revendication 49 caractérisé en ce qu'il s'agit d'un rétrovirus recombinant défectif.

52. Composition pharmaceutique comprenant au moins un vecteur selon l'une des revendications 41 à 51.

53. Acide nucléique comprenant la séquence SEQ ID n° 4.

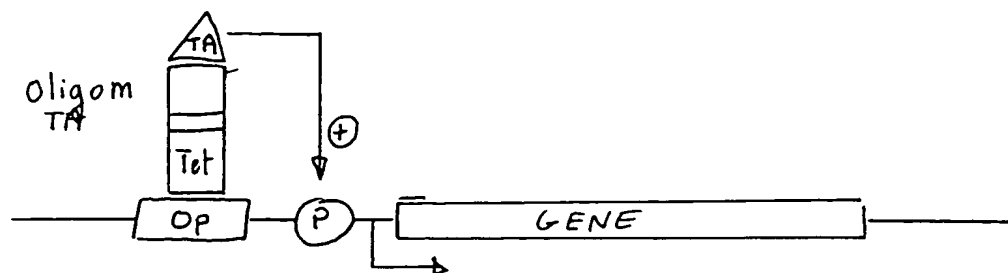


Figure 1A

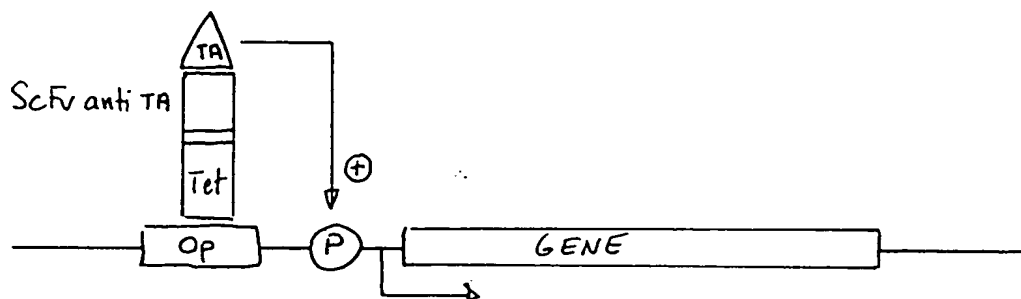


Figure 1B

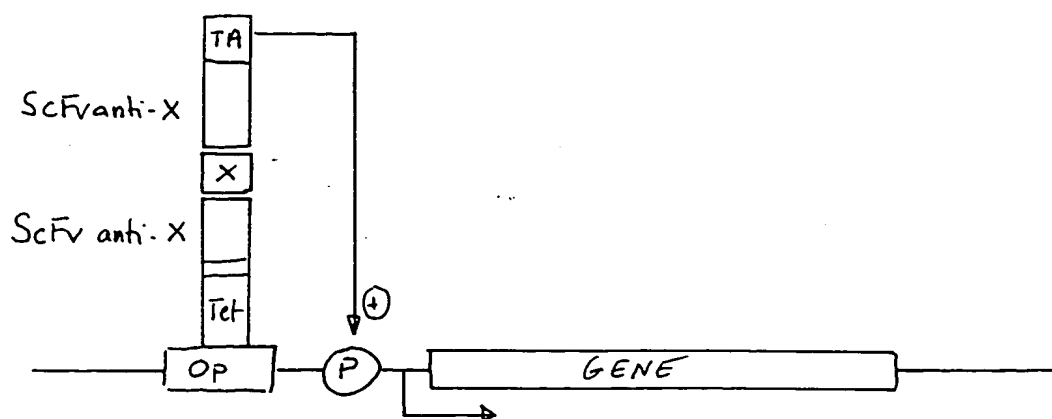
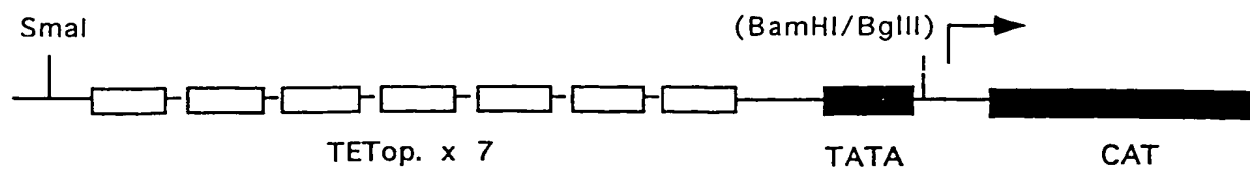


Figure 2



TETOp. = GACTTTCAC TTTTCTCTATCACTGATAGGGAGTGGTAAACTC

Figure 3

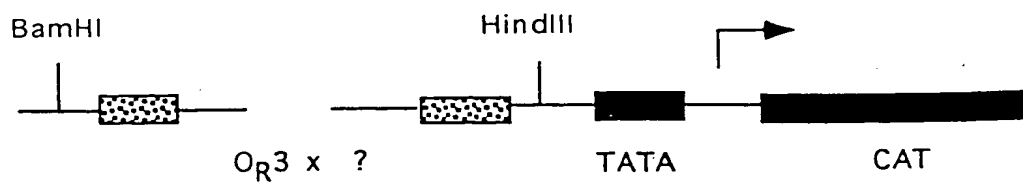


Figure 4

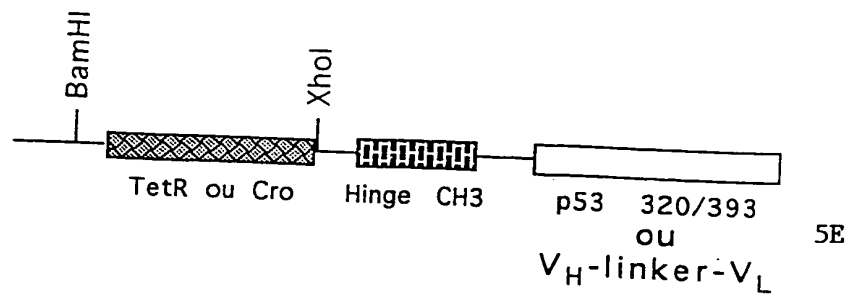
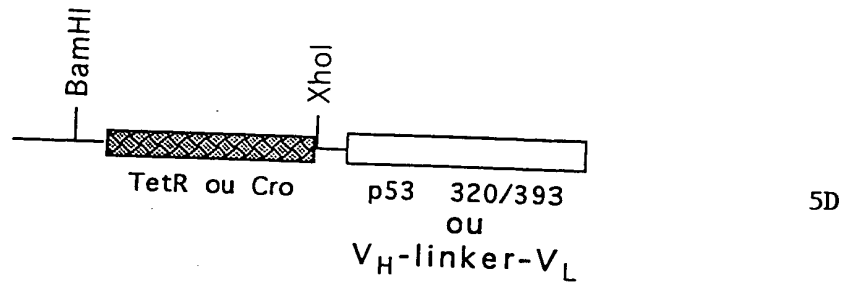
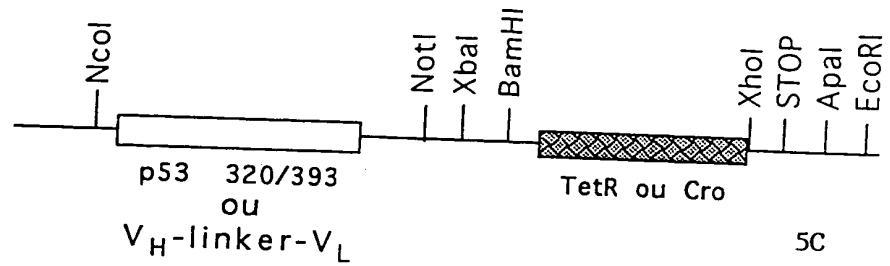
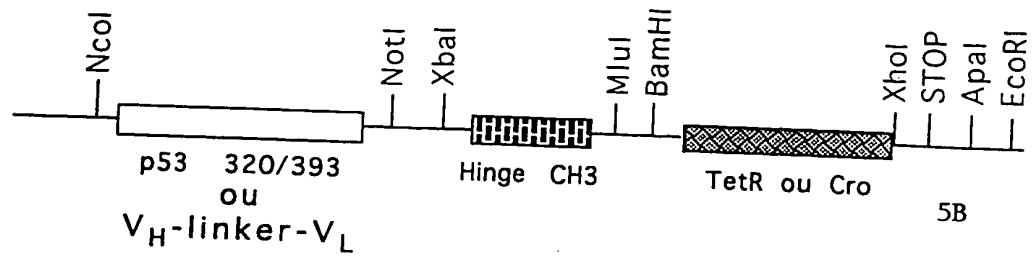
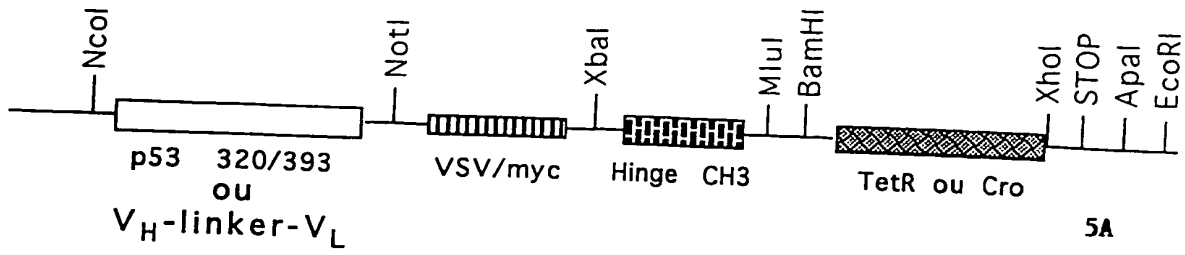
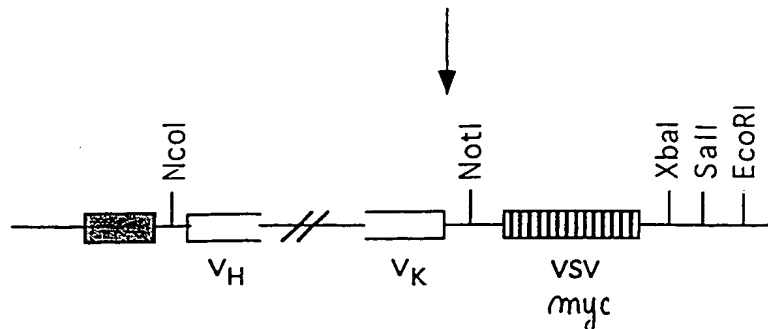


Figure 5



TET Rep.
XhoI
Apal
EcoRI

... CAA CTT AAA TGT GAA AGT GGG TCC CTC GAG TAA GGG CCC GAA TTC CGT ACG

(5475) CGT ACG GAA TTC GGG CCC TTA CTC GAG GGA CCC ACT TTC ACA TTT AAG TTG

(5531) XbaI Hinge CH3 MluI BamHI  
GGC TCT AGA CCC AAG CCC AGT ACC CCC CCA GGT TCT TCA ACG CGT GGA TCC ATG GAA CAA CGC ATA ACC CTG AAA G  
ATG GAA CAA CGC ATA ACC CTG AAA G  
CRO XhoI Apal EcoRI CRO  
... CCG AGT AAC AAA AAA ACA ACA GCA CTC GAG TAA GGG CCC GAA TTC CGT ACG

(5532) CGT ACG GAA TTC GGG CCC TTA CTC GAG TGC TGT TGT TTT TTT GTT ACT CGG

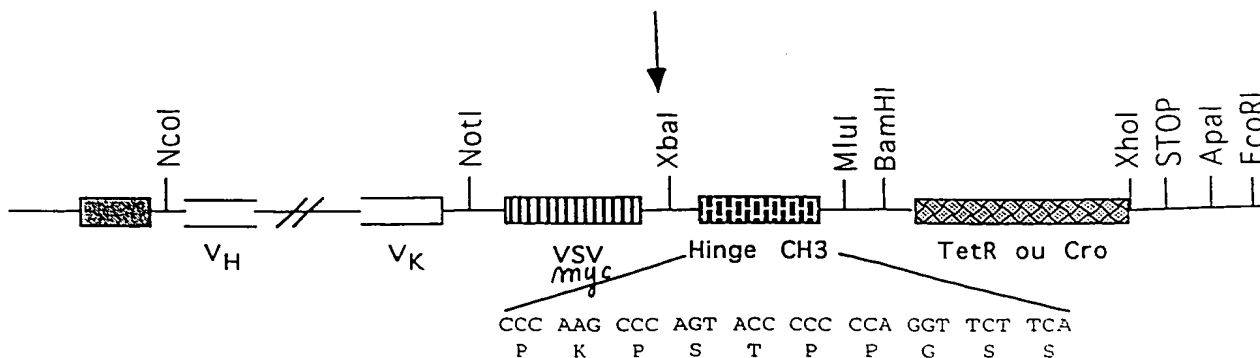


Figure 6

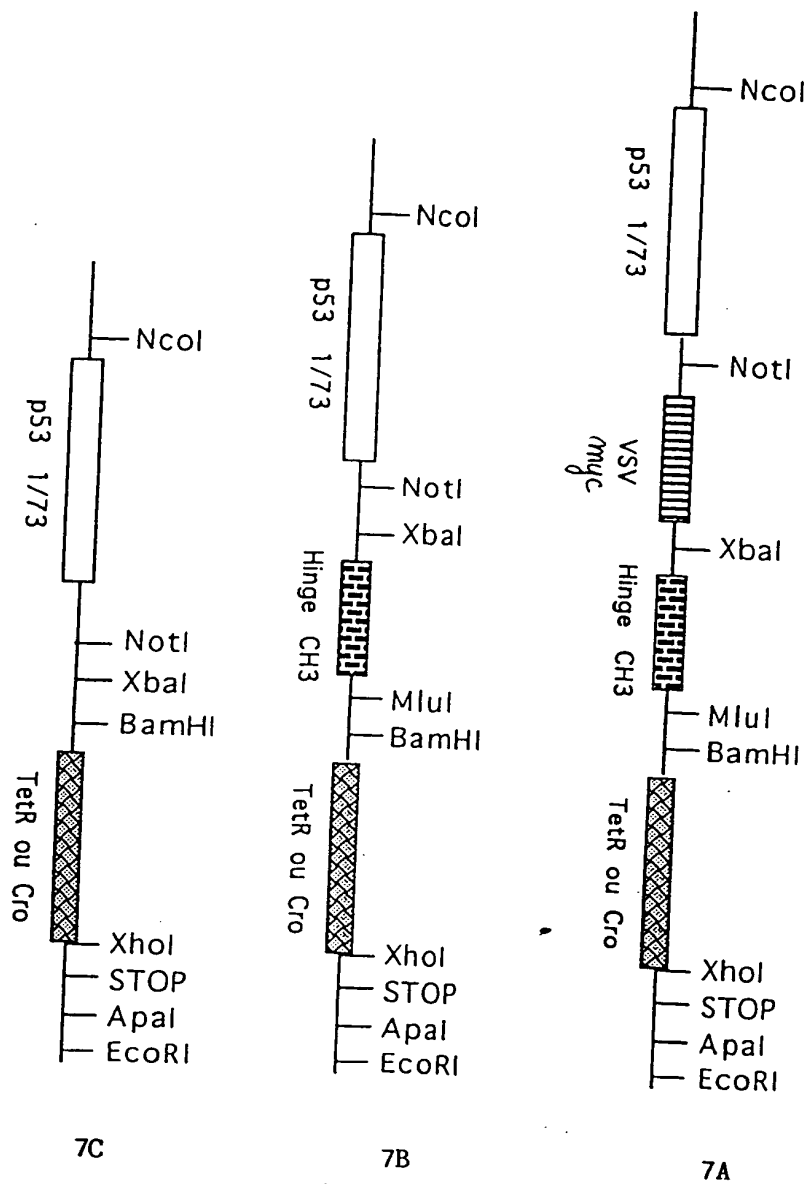


Figure 7



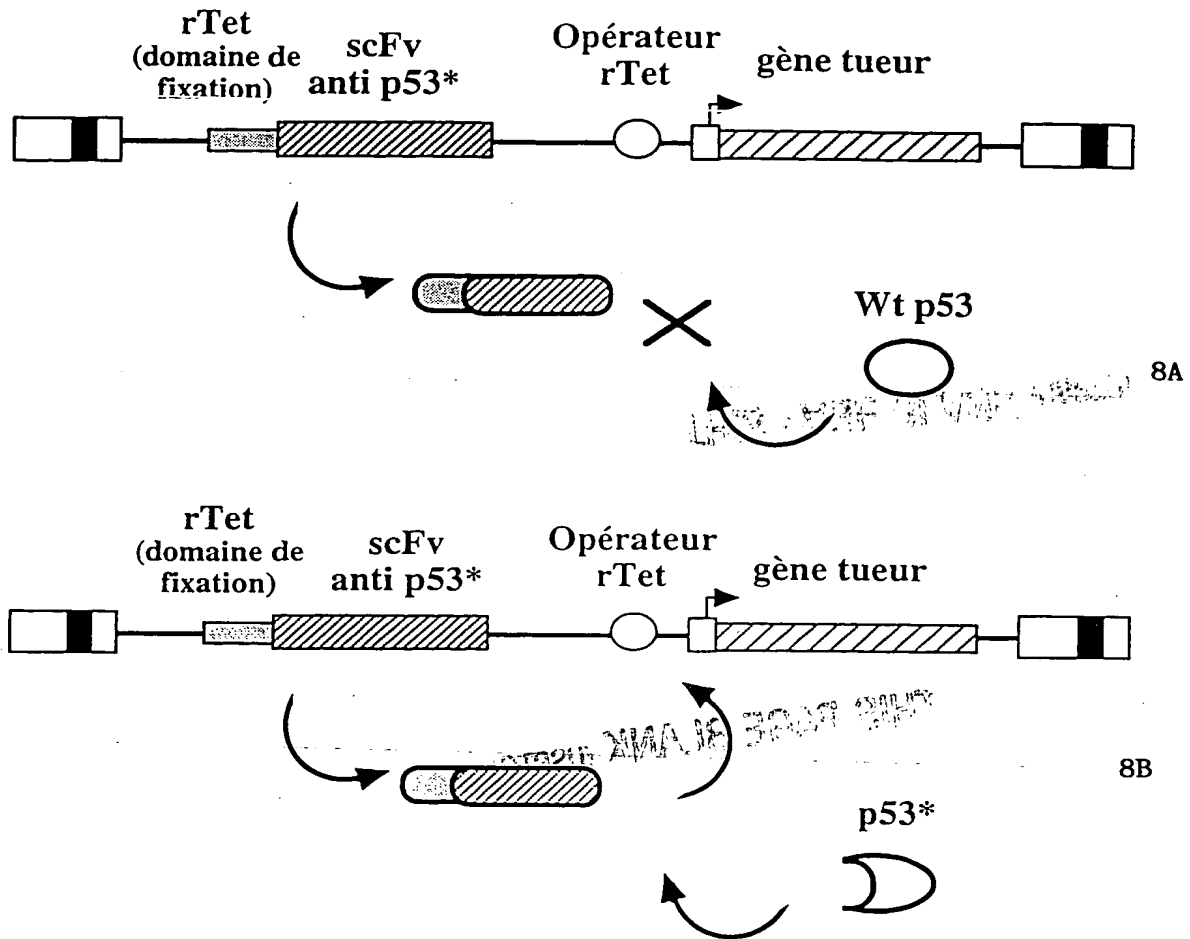


Figure 8

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**